

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med* 2011;364:2496-506.

Supplementary Methods

Single Nucleotide Polymorphism (SNP) Array. DNA from 75 MDS patient samples was prepared and hybridized to Affymetrix Genome-Wide Human SNP 6.0 Array GeneChip microarrays according to the manufacturer's protocols (Affymetrix, Santa Clara, CA). Copy number variants were detected using the Ulrasome aberration caller (Nilsson et al. Ulrasome: efficient aberration caller for copy number studies of ultra-high resolution. *Bioinformatics* 2009;25:1078-9)

454 DNA Sequencing and Analysis. Pools of 4 or 5 samples were combined prior to PCR amplification of target regions. Candidate mutations were called using the GS Amplicon Variant Analyzer Software and filtered for a frequency of 2% or greater. Candidate insertions and single base deletions were called by analyzing each individual aligned read and removing variants present at less than 2% frequency or predominantly in one direction. All candidate mutations were manually reviewed to confirm alignment. Mutations discovered in bidirectional reads with at least 5 reads in one direction were selected for validation by hME or Sanger sequencing.

Sanger DNA Sequencing and Analysis. Target regions in individual patient samples were PCR amplified techniques and sequenced using standard techniques, yielding 95.9% of all trimmed reads with an average quality score of 20 or more. Chromatogram data was analyzed with Mutation Surveyor v3.30 (State College, PA). All traces were visually inspected to identify and confirm mutation candidates.

Statistical Analysis. Patient characteristics were compared between groups using the Fishers exact test for categorical data, the Jonckheere-Terpstra test for ordered categorical data, and the Wilcoxon rank-sum test for continuous data (Conover WJ. *Practical nonparametric statistics*. 3rd ed. New York: Wiley; 1999). Overall survival (OS) was measured from the time of sample collection to time of death from any cause; patients last known to be alive were censored at that time. OS curves were constructed using the method of Kaplan and Meier and compared using the log-rank test. All *P* values were based on 2-sided tests. For the univariate analyses of the association of clinical characteristics with each of the 18 mutations a *P* ≤ 0.01 was considered statistically significant, to reflect the multiplicity of clinical features of the individual patient. For all other assessments, nominal *p*-values are presented.

OS was evaluated for all patients using unadjusted and adjusted Cox proportional hazard regression modeling; models were adjusted for IPSS risk group at the time of sample collection. The prognostic significance of each mutation was determined using step-up models evaluated using the -2 Log likelihood statistic (Collett D. *Modelling survival data in medical research*. 2nd ed. Boca Raton, Fla.: Chapman & Hall/CRC; 2003). The full model including mutational status and IPSS risk was compared to the null model including only IPSS risk using the difference in -2 Log likelihood and tested using a chi-square distribution with the appropriate degrees of freedom. Candidate explanatory variables in the stepwise Cox regression modeling included age (categorized as <55 vs. ≥55 years), IPSS risk, sex, and the 13 mutations with ≥1% frequency (of the 18 examined as potential prognostic features). The same final model was obtained using a forward variable selection procedure. The categories for age were determined using a recursive partitioning algorithm based on OS. SAS version 9.2 and R version 2.8.0 were used for all analyses.

Supplementary Tables

Supplementary Table 1: Patient Characteristics and Association with Median Survival

	N (%)	Median Survival years (95% CI)	p-value†
Age			
< 55 yrs.	49 (11)	3.14 (2.54, 7.44)	0.003
55 - 64 yrs.	88 (20)	2.02 (1.29, 2.72)	
65 - 74 yrs.	179 (41)	1.83 (1.47, 2.14)	
≥ 75 yrs.	123 (28)	1.44 (1.05, 1.86)	
Sex			
Female	133 (30)	2.14 (1.71, 2.83)	0.10
Male	306 (70)	1.72 (1.40, 2.03)	
FAB			
RA	197 (45)	2.62 (2.13, 3.12)	<0.001
RARS	47 (11)	4.16 (1.60, 5.19)	
RAEB-I	105 (24)	1.40 (1.07, 1.86)	
RAEB-II	55 (13)	0.95 (0.76, 1.22)	
RAEB-t	34 (8)	1.11 (0.61, 1.47)	
Unknown	1 (<1)	NA	
IPSS			
Low	110 (25)	4.23 (3.14, 5.55)	<0.001
Int1	185 (42)	1.86 (1.61, 2.24)	
Int2	101 (23)	0.95 (0.73, 1.22)	
High	32 (7)	0.79 (0.48, 1.11)	
Unknown	11 (3)	NA	
Karyotype			
Good	310 (71)	2.18 (1.86, 2.59)	<0.001
Intermediate	55 (13)	1.37 (0.84, 1.86)	
Poor	67 (15)	0.87 (0.58, 1.27)	
Unknown	7 (1)	NA	
Karyotype			
-7/del(7q) isolated or +1	10 (2)	0.73 (0.07, 1.05)	<0.001
del(20q) isolated	18 (4)	1.32 (0.88, 1.98)	
del(5q) isolated	22 (5)	1.67 (1.02, 4.16)	
+8 isolated	24 (5)	1.40 (0.79, 2.14)	
Complex	57 (13)	1.10 (0.58, 1.70)	
Normal	255 (58)	2.40 (2.11, 2.72)	
Other	46 (10)	1.31 (0.78, 2.11)	
Unknown	7 (1)	NA	
Blast %			
< 5 %	247 (56)	2.65 (2.24, 3.14)	<0.001
5–10 %	112 (26)	1.34 (0.96, 1.71)	

11–20 %	61 (14)	1.06 (0.77, 1.26)	
21–30 %	18 (4)	1.15 (0.61, 1.48)	
Unknown	1 (<1)	NA	
Hemoglobin (g/dL)			
< 8.0	46 (10)	1.22 (0.70, 1.98)	0.002
8.0 – 9.99	175 (40)	1.51 (1.11, 1.88)	
10.0 – 11.99	145 (33)	2.14 (1.70, 2.59)	
≥12.0	64 (15)	3.05 (2.03, 6.28)	
Unknown	9 (2)	NA	
Absolute Neutrophil Count (cells/ μ L)			
< 500	58 (13)	1.47 (0.88, 1.93)	0.001
500 – 1,499	145 (33)	1.59 (1.11, 2.14)	
1,500 – 9,999	203 (46)	2.54 (2.02, 2.89)	
≥10,000	8 (2)	0.91 (0.14, 3.33)	
Unknown	25 (6)	NA	
Platelets (per μ L)			
< 50,000	123 (28)	1.07 (0.79, 1.36)	<0.001
50,000 – 149,000	168 (38)	1.82 (1.40, 2.14)	
150,000 – 449,999	119 (27)	3.18 (2.54, 4.89)	
≥ 450,000	20 (5)	3.73 (1.46, 8.65)	
Unknown	9 (2)	NA	

[†]Testing excludes the unknown groups. Median survival time estimates are not provided for the unknown groups.

A total of 31 patients were reported to have therapy-related MDS. Since the IPSS has not been validated in this subgroup, and it is difficult to be certain which patients with an exposure history truly have therapy-related MDS, we assigned an IPSS risk group to all patients.

Only 10 patients in our sample set underwent stem-cell transplantation.

Details of other potential treatments are not known for all patients. However, a large proportion of patients died before the FDA approval of therapies proven to extend overall survival of patients with MDS. Of the 332 patients that are known to have died during the follow up period, 167 (50.3%) died prior to the date of azacitidine approval (5/19/2004), 201 (63.3%) died prior to the date of decitabine approval (5/2/2006), and 193 (58.1%) died prior to the date of lenalidomide approval (12/27/2005). Therefore, many of patients examined in this study are unlikely to have received potentially life prolonging treatment.

Supplementary Table 2: Determination of International Prognostic Scoring System (IPSS) Risk Groups.

Each of three prognostic variables is assigned a score as outlined in the table below. The sum of these scores is used to determine the IPSS risk group.

International Prognostic Scoring System					
Prognostic Variable	Score Value				
	0	0.5	1.0	1.5	2.0
Bone Marrow Blast Percentage (%)	< 5	5-10		11-20	21-30
Karyotype Category*	Good	Intermediate	Poor		
Number of Cytopenias**	0 or 1	2 or 3			

*Karyotype Categories
Good: normal, -Y, del(5q), del(20q)
Intermediate: other (not good or poor)
Poor: chromosome 7 anomalies or complex (≥ 3 abnormalities)

**Cytopenia Definitions
Hemoglobin: < 10 g/dL
Absolute Neutrophil Count: < 1800 per μ L
Platelet Count: < 100,000 per μ L

IPSS Risk Groups	Total Score
Low	0
Intermediate-1	0.5-1.0
Intermediate-2	1.5-2.0
High	≥ 2.5

Tables adapted from:

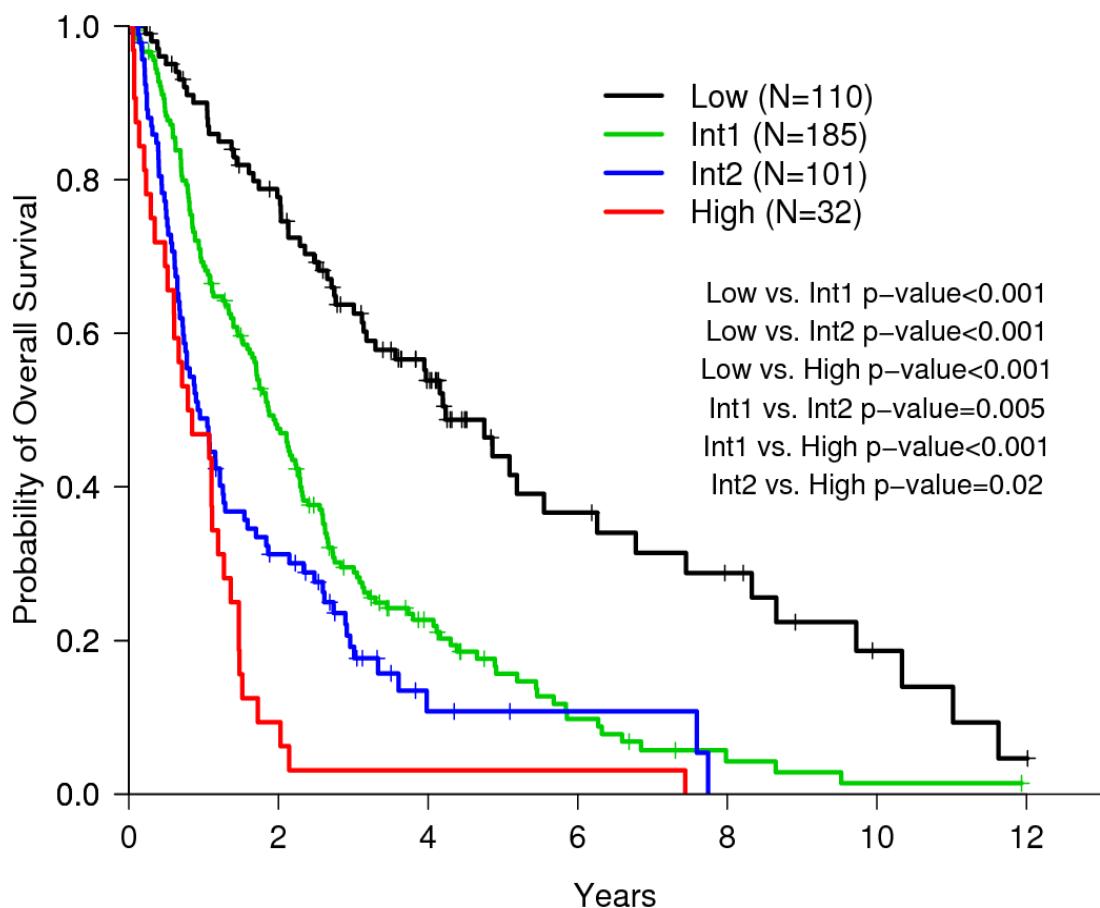
Greenberg P, Cox C, LeBeau MM, et al. International Scoring System for Evaluating Prognosis in Myelodysplastic Syndromes. *Blood* 1997;89:2079-88.

Supplementary Table 3: IPSS Classification and Association with Overall Survival. IPSS risk group was determined at the time of sample collection and compared to the original IPSS classification for each patient. As shown in Table A, 73 patients had their IPSS risk group change upon reclassification. Only 2 patients moved more than one category (1 from Int2 to Low and 1 from Low to Int2). Figure B is a Kaplan-Meier survival plot comparing the overall survival of patients based on their reclassified IPSS. Pairwise comparisons confirm that survival of each group is significantly different from every other group.

A.

		IPSS Classification at Time of Sample Collection				
Initial IPSS Classification	N	Low	Int1	Int2	High	Unknown
	Low	86	11	1	0	0
	Int1	22	160	8	0	2
	Int2	1	10	88	2	3
	High	0	0	4	30	4
	Unknown	1	4	0	0	2

B.



Supplementary Table 4: List of mutations predicted to disrupt protein sequence and included in the analysis.

Gene Name	Ensembl Reference Transcript	RefSeq Reference Transcript	DNA Mutation	Protein Mutation	Confirmed Somatic in One or More Samples	Present in Two or More Samples
ASXL1	ENST00000375687	NM_015338.5	c.1205G>A	p.R402Q	YES	
ASXL1	ENST00000375687	NM_015338.5	c.1249C>T	p.R417*	YES	
ASXL1	ENST00000375687	NM_015338.5	c.1282C>T	p.Q428*		
ASXL1	ENST00000375687	NM_015338.5	c.1285G>T	p.E429*		
ASXL1	ENST00000375687	NM_015338.5	c.1291G>T	p.E431*	YES	
ASXL1	ENST00000375687	NM_015338.5	c.1534C>T	p.Q512*		
ASXL1	ENST00000375687	NM_015338.5	c.1747_1757delTGGGTGGTTAA	p.W583RfsX32		
ASXL1	ENST00000375687	NM_015338.5	c.1772dupA	p.Y591*		YES
ASXL1	ENST00000375687	NM_015338.5	c.1900_1922delAGAGAGGGGCCACCTGCCAT	p.E635RfsX15	YES	YES
ASXL1	ENST00000375687	NM_015338.5	c.1917_1928delTGCCATCGGAGGinsC	p.A640GfsX14	YES	
ASXL1	ENST00000375687	NM_015338.5	c.1924_1929delGGAGGGinsT	p.G642WfsX14	YES	
ASXL1	ENST00000375687	NM_015338.5	c.1926_1930delAGGGG	p.G644WfsX12		
ASXL1	ENST00000375687	NM_015338.5	c.1934dupG	p.G646WfsX12	YES	
ASXL1	ENST00000375687	NM_015338.5	c.1979_1982dupGCAG	p.G662QfsX7		
ASXL1	ENST00000375687	NM_015338.5	c.2035G>T	p.G679*		
ASXL1	ENST00000375687	NM_015338.5	c.2036dupG	p.G680RfsX38		
ASXL1	ENST00000375687	NM_015338.5	c.2041_2042delCC	p.P681EfsX36		
ASXL1	ENST00000375687	NM_015338.5	c.2069_2076delATCTACAG	p.D690AfsX25	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2077C>T	p.R693*	YES	YES
ASXL1	ENST00000375687	NM_015338.5	c.2088_2111delACTGCCGCTTATCCTCTAAATGGinsTAGA	p.L697RfsX14		
ASXL1	ENST00000375687	NM_015338.5	c.2110G>A	p.G704R		
ASXL1	ENST00000375687	NM_015338.5	c.2172_2176dupGAGAA	p.K726RfsX20	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2226delG	p.L743*		
ASXL1	ENST00000375687	NM_015338.5	c.2242C>T	p.Q748*		
ASXL1	ENST00000375687	NM_015338.5	c.2253delT	p.A752LfsX20		
ASXL1	ENST00000375687	NM_015338.5	c.2269C>T	p.Q757*	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2278C>T	p.Q760*	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2322delA	p.R774SfsX2		
ASXL1	ENST00000375687	NM_015338.5	c.2332C>T	p.Q778*		
ASXL1	ENST00000375687	NM_015338.5	c.2395G>T	p.D799Y		
ASXL1	ENST00000375687	NM_015338.5	c.2415dupC	p.T806HfsX16	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2423dupC	p.A809CfsX13		
ASXL1	ENST00000375687	NM_015338.5	c.2485C>T	p.Q829*		
ASXL1	ENST00000375687	NM_015338.5	c.2629G>T	p.E877*		
ASXL1	ENST00000375687	NM_015338.5	c.2681delG	p.S894IfsX14		
ASXL1	ENST00000375687	NM_015338.5	c.2693G>A	p.W898*		
ASXL1	ENST00000375687	NM_015338.5	c.2708C>A	p.S903*		
ASXL1	ENST00000375687	NM_015338.5	c.2757dupA	p.P920TfsX4		
ASXL1	ENST00000375687	NM_015338.5	c.2777_2780dupTTGG	p.E928WfsX21	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2917dupA	p.S973KfsX9	YES	
ASXL1	ENST00000375687	NM_015338.5	c.2945delA	p.K982SfsX2		
ASXL1	ENST00000375687	NM_015338.5	c.2973_2985delACTGAGTCCTCAC	p.L992VfsX28		
ASXL1	ENST00000375687	NM_015338.5	c.3050delA	p.D1017AfsX7		
ASXL1	ENST00000375687	NM_015338.5	c.3083C>A	p.S1028*		
ASXL1	ENST00000375687	NM_015338.5	c.3187C>T	p.Q1063*		
ASXL1	ENST00000375687	NM_015338.5	c.3401C>T	p.P1134L		
ASXL1	ENST00000375687	NM_015338.5	c.3601delC	p.Q1201KfsX16	YES	
ASXL1	ENST00000375687	NM_015338.5	c.3758_3759insC	p.N1254*		
ASXL1	ENST00000375687	NM_015338.5	c.4456G>A	p.A1486T	YES	
BRAF	ENST00000288602	NM_004333.4	c.1790T>A	p.L597Q		
BRAF	ENST00000288602	NM_004333.4	c.1799T>A	p.V600E		
CBL	ENST00000264033	NM_005188.2	c.1122_1127delGGGCTC	p.M374_S376delinsI		
CBL	ENST00000264033	NM_005188.2	c.1142G>A	p.C381Y	YES	
CBL	ENST00000264033	NM_005188.2	c.1143T>G	p.C381W	YES	
CBL	ENST00000264033	NM_005188.2	c.1151G>A	p.C384Y		
CBL	ENST00000264033	NM_005188.2	c.1216A>C	p.T406P		
CBL	ENST00000264033	NM_005188.2	c.1247G>A	p.C416Y	YES	
CBL	ENST00000264033	NM_005188.2	c.1253T>C	p.F418S	YES	
CBL	ENST00000264033	NM_005188.2	c.1258C>T	p.R420*	YES	
CBL	ENST00000264033	NM_005188.2	c.1259G>A	p.R420Q		YES
CBL	ENST00000264033	NM_005188.2	c.1292T>C	p.V431A	YES	
CBL	ENST00000264033	NM_005188.2	c.1301T>C	p.F434S		
CDKN2A - p16INK4A / p14ARF	ENST00000304494 / ENST00000361570	NM_000077.3 / NM_058195.2	c.198C>G / c.364C>G	p.H66Q / p.R122G	YES	
ELANE	ENST00000263621	NM_001972.2	c.257C>T	p.A86V		
ELANE	ENST00000263621	NM_001972.2	c.751G>T	p.D251Y		
ETV6	ENST00000266427	NM_001987.4	c.195_197dupCGT	p.A66_V67insV		
ETV6	ENST00000266427	NM_001987.4	c.290T>C	p.L97P		
ETV6	ENST00000266427	NM_001987.4	c.306dupT	p.R103SfsX9		
ETV6	ENST00000266427	NM_001987.4	c.380G>A	p.R127Q		

ETV6	ENST00000266427	NM_001987.4	c.419_420dupTA	p.H141YfsX69		
ETV6	ENST00000266427	NM_001987.4	c.463G>C	p.D155H	YES	
ETV6	ENST00000266427	NM_001987.4	c.537dupG	p.H180AfsX16	YES	
ETV6	ENST00000266427	NM_001987.4	c.602T>C	p.L201P	YES	YES
ETV6	ENST00000266427	NM_001987.4	c.1075C>T	p.R359*		
EZH2	ENST00000320356	NM_004456.3	c.72dupG	p.R25AfsX12		
EZH2	ENST00000320356	NM_004456.3	c.130delT	p.S44PfsX13		
EZH2	ENST00000320356	NM_004456.3	c.151G>T	p.E51*		
EZH2	ENST00000320356	NM_004456.3	c.187C>T	p.R63*	YES	
EZH2	ENST00000320356	NM_004456.3	c.349C>T	p.Q117*	YES	
EZH2	ENST00000320356	NM_004456.3	c.393_394delTCinsGA	p.I131_P132delinsMT		
EZH2	ENST00000320356	NM_004456.3	c.401T>A	p.M134K	YES	
EZH2	ENST00000320356	NM_004456.3	c.487_507delTGCGGTTATAATGATGAA	p.I63_169delCGFINDE		
EZH2	ENST00000320356	NM_004456.3	c.698A>G	p.D233G		
EZH2	ENST00000320356	NM_004456.3	c.729_2A>T	Splice Disruption	YES	
EZH2	ENST00000320356	NM_004456.3	c.893G>A	p.R298H		
EZH2	ENST00000320356	NM_004456.3	c.1119dupC	p.T374HfsX3		
EZH2	ENST00000320356	NM_004456.3	c.1370G>A	p.C457Y		
EZH2	ENST00000320356	NM_004456.3	c.1483delA	p.R495GfsX19		
EZH2	ENST00000320356	NM_004456.3	c.1505+5G>A	Splice Disruption		
EZH2	ENST00000320356	NM_004456.3	c.1615T>C	p.C539R	YES	
EZH2	ENST00000320356	NM_004456.3	c.1700G>A	p.C567Y		
EZH2	ENST00000320356	NM_004456.3	c.1723C>T	p.Q575*	YES	
EZH2	ENST00000320356	NM_004456.3	c.1739T>C	p.L580P	YES	
EZH2	ENST00000320356	NM_004456.3	c.1747C>T	p.R583*	YES	
EZH2	ENST00000320356	NM_004456.3	c.1852_6C>T	Splice Disruption		
EZH2	ENST00000320356	NM_004456.3	c.1957C>G	p.Q653E	YES	
EZH2	ENST00000320356	NM_004456.3	c.1969G>C	p.D657H	YES	YES
EZH2	ENST00000320356	NM_004456.3	c.1987T>C	p.Y663H		
EZH2	ENST00000320356	NM_004456.3	c.2008T>C	p.F670L	YES	
EZH2	ENST00000320356	NM_004456.3	c.2029+1G>T	Splice Disruption		
EZH2	ENST00000320356	NM_004456.3	c.2069G>A	p.R690H	YES	
EZH2	ENST00000320356	NM_004456.3	c.2110+1G>T	Splice Disruption		
EZH2	ENST00000320356	NM_004456.3	c.2111_2A>T	Splice Disruption		
EZH2	ENST00000320356	NM_004456.3	c.2187delT	p.F729LfsX11		
FLNB	ENST00000295956	NM_001457.3	c.1697G>A	p.R566Q		
GNAS	ENST00000371082	NM_000516.4	c.601C>T	p.R201C		
IDH1	ENST00000415913	NM_005896.2	c.394C>G	p.R201H	YES	YES
IDH1	ENST00000415913	NM_005896.2	c.394C>T	p.R132G		
IDH1	ENST00000415913	NM_005896.2	c.395G>A	p.R132H	YES	
IDH2	ENST00000330062	NM_002168.2	c.419G>A	p.R140Q	YES	YES
IDH2	ENST00000330062	NM_002168.2	c.515G>A	p.R172K	YES	
JAK2	ENST00000381652	NM_004972.3	c.1849G>T	p.V617F	YES	YES
KRAS	ENST00000395977	NM_033360.2	c.34G>C	p.G12R	YES	
KRAS	ENST00000395977	NM_033360.2	c.35G>T	p.G12V	YES	
KRAS	ENST00000395977	NM_033360.2	c.183C>T	p.Q61H		
KRAS	ENST00000395977	NM_033360.2	c.436G>A	p.A146T		
NPM1	ENST00000296930	NM_002520.5	c.860_863dupTCTG	p.W288CfsX12	YES	
NPM1	ENST00000296930	NM_002520.5	c.863_864insCATG	p.W288CfsX12		
NPM1	ENST00000296930	NM_002520.5	c.863_864insTATG	p.W288CfsX12		
NRAS	ENST00000369535	NM_002524.3	c.34G>A	p.G12S	YES	YES
NRAS	ENST00000369535	NM_002524.3	c.34G>C	p.G12R	YES	
NRAS	ENST00000369535	NM_002524.3	c.35G>A	p.G12D	YES	
NRAS	ENST00000369535	NM_002524.3	c.35G>T	p.G12V	YES	
NRAS	ENST00000369535	NM_002524.3	c.37G>C	p.G13R		
NRAS	ENST00000369535	NM_002524.3	c.38G>A	p.G13D	YES	
NRAS	ENST00000369535	NM_002524.3	c.38G>T	p.G13V		YES
NRAS	ENST00000369535	NM_002524.3	c.181C>A	p.Q61K		
PTEN	ENST00000371953	NM_000314.4	c.947T>A	p.L316Q	YES	
PTPN11	ENST00000351677	NM_002834.3	c.214G>A	p.A72T	YES	YES
PTPN11	ENST00000351677	NM_002834.3	c.226G>A	p.E76K	YES	
RUNX1	ENST00000300305	NM_001754.4	c.352_1G>A	Splice Disruption	YES	
RUNX1	ENST00000300305	NM_001754.4	c.364G>A	p.G122R		
RUNX1	ENST00000300305	NM_001754.4	c.398_400dupTGG	p.M133_A134insV		
RUNX1	ENST00000300305	NM_001754.4	c.401C>T	p.A134V		
RUNX1	ENST00000300305	NM_001754.4	c.404G>A	p.G135D	YES	
RUNX1	ENST00000300305	NM_001754.4	c.422C>T	p.S141L	YES	YES
RUNX1	ENST00000300305	NM_001754.4	c.436A>G	p.N146D		
RUNX1	ENST00000300305	NM_001754.4	c.432_441dupGAGAAATGCT	p.T148EfsX15		
RUNX1	ENST00000300305	NM_001754.4	c.472T>A	p.F158I	YES	
RUNX1	ENST00000300305	NM_001754.4	c.479A>G	p.D160G		
RUNX1	ENST00000300305	NM_001754.4	c.482T>C	p.L161P	YES	
RUNX1	ENST00000300305	NM_001754.4	c.486G>C	p.R162S		
RUNX1	ENST00000300305	NM_001754.4	c.496C>G	p.R166G	YES	
RUNX1	ENST00000300305	NM_001754.4	c.496C>T	p.R166*		
RUNX1	ENST00000300305	NM_001754.4	c.497G>A	p.R166Q		
RUNX1	ENST00000300305	NM_001754.4	c.504_508+1dupAAGAGG	p.G168_R169dup		
RUNX1	ENST00000300305	NM_001754.4	c.494_497dupGTG	p.G168KfsX46		
RUNX1	ENST00000300305	NM_001754.4	c.5084_508+5insCAAGGAAAAA	Splice Disruption		
RUNX1	ENST00000300305	NM_001754.4	c.508+5G>A	Splice Disruption		
RUNX1	ENST00000300305	NM_001754.4	c.511A>T	p.K171*		

RUNX1	ENST00000300305	NM_001754.4	c.592G>A		p.D198N		YES
RUNX1	ENST00000300305	NM_001754.4	c.601C>T		p.R201*	YES	YES
RUNX1	ENST00000300305	NM_001754.4	c.602G>A		p.R201Q		
RUNX1	ENST00000300305	NM_001754.4	c.610C>T		p.R204*	YES	
RUNX1	ENST00000300305	NM_001754.4	c.613+2T>G		Splice Disruption	YES	
RUNX1	ENST00000300305	NM_001754.4	c.620G>A		p.R207Q		
RUNX1	ENST00000300305	NM_001754.4	c.714_727delCAGCCCCACACCAC		p.H242AfsX14	YES	
RUNX1	ENST00000300305	NM_001754.4	c.861C>A		p.Y287*	YES	
RUNX1	ENST00000300305	NM_001754.4	c.961_962delCT		p.S322NfsX277	YES	
RUNX1	ENST00000300305	NM_001754.4	c.965C>G		p.S322*	YES	
RUNX1	ENST00000300305	NM_001754.4	c.1057_1070delTTTCACCTACTCCCC		p.F53DfsX242		
RUNX1	ENST00000300305	NM_001754.4	c.1070delC		p.P357RfsX237		
RUNX1	ENST00000300305	NM_001754.4	c.1203dupC		p.S402LfsX198		
RUNX1	ENST00000300305	NM_001754.4	c.1203_1206dupCTCC		p.Y403LfsX198		
RUNX1	ENST00000300305	NM_001754.4	c.1208_1209insT		p.H404PfsX196		
RUNX1	ENST00000300305	NM_001754.4	c.1239_1246delCTACCACT		p.Y414LfsX183	YES	
RUNX1	ENST00000300305	NM_001754.4	c.1274C>T		p.P425L		
TET2	ENST00000380013	NM_001127208.1	c.428delA		p.D143VfsX2		
TET2	ENST00000380013	NM_001127208.1	c.444_468dupAGAACCTGTGAGTTCTGTAGCCCCAA		p.E157RfsX13		
TET2	ENST00000380013	NM_001127208.1	c.551_552delAG		p.E184AfsX7		
TET2	ENST00000380013	NM_001127208.1	c.578delA		p.H193LfsX14		
TET2	ENST00000380013	NM_001127208.1	c.651delC		p.V218WfsX32		
TET2	ENST00000380013	NM_001127208.1	c.727C>T		p.Q243*		
TET2	ENST00000380013	NM_001127208.1	c.744C>A		p.H248Q	YES	
TET2	ENST00000380013	NM_001127208.1	c.822delC		p.N275IfsX18		YES
TET2	ENST00000380013	NM_001127208.1	c.1122delT		p.N374KfsX3		
TET2	ENST00000380013	NM_001127208.1	c.1128G>A		p.M376I		
TET2	ENST00000380013	NM_001127208.1	c.1128delG		p.M376IfsX51		
TET2	ENST00000380013	NM_001127208.1	c.1326_1327delCA		p.T443NfsX11		
TET2	ENST00000380013	NM_001127208.1	c.1342dupG		p.E448GfsX7		
TET2	ENST00000380013	NM_001127208.1	c.1473_1497delGACTGCAGGGACAATGACTGTTCCA		p.Q491HfsX34	YES	YES
TET2	ENST00000380013	NM_001127208.1	c.1591C>T		p.Q531*	YES	
TET2	ENST00000380013	NM_001127208.1	c.1630C>T		p.R544*		
TET2	ENST00000380013	NM_001127208.1	c.1648C>T		p.R550*		
TET2	ENST00000380013	NM_001127208.1	c.1699_1703delTTGAA		p.L567GfsX14	YES	
TET2	ENST00000380013	NM_001127208.1	c.1795C>T		p.Q599*	YES	
TET2	ENST00000380013	NM_001127208.1	c.1903C>T		p.Q635*		
TET2	ENST00000380013	NM_001127208.1	c.1924C>T		p.Q642*		
TET2	ENST00000380013	NM_001127208.1	c.1928_1935delCCCAAGGT		p.S643YfsX35		
TET2	ENST00000380013	NM_001127208.1	c.1970C>G		p.S657*	YES	
TET2	ENST00000380013	NM_001127208.1	c.2053C>T		p.Q685*		YES
TET2	ENST00000380013	NM_001127208.1	c.2101C>T		p.Q701*	YES	
TET2	ENST00000380013	NM_001127208.1	c.2176C>T		p.Q726*		
TET2	ENST00000380013	NM_001127208.1	c.2233C>T		p.Q745*	YES	
TET2	ENST00000380013	NM_001127208.1	c.2249_2252delTAAA		p.T750RfsX62		
TET2	ENST00000380013	NM_001127208.1	c.2280delT		p.P761LfsX52		
TET2	ENST00000380013	NM_001127208.1	c.2318_2321dupGATC		p.F775IfsX7		
TET2	ENST00000380013	NM_001127208.1	c.2386G>C		p.E796Q	YES	
TET2	ENST00000380013	NM_001127208.1	c.2392G>T		p.E798*		
TET2	ENST00000380013	NM_001127208.1	c.2525C>G		p.S842*	YES	
TET2	ENST00000380013	NM_001127208.1	c.2554G>T		p.E852*		
TET2	ENST00000380013	NM_001127208.1	c.2650C>T		p.Q884*		
TET2	ENST00000380013	NM_001127208.1	c.2759dupT		p.L920FfsX4		
TET2	ENST00000380013	NM_001127208.1	c.3056delT		p.V1019GfsX14		
TET2	ENST00000380013	NM_001127208.1	c.3130A>T		p.K1044*	YES	
TET2	ENST00000380013	NM_001127208.1	c.3245_3246delAG		p.E1082AfsX21	YES	
TET2	ENST00000380013	NM_001127208.1	c.3311_3315delTTATA		p.F1104*		
TET2	ENST00000380013	NM_001127208.1	c.3316G>T		p.E1106*		
TET2	ENST00000380013	NM_001127208.1	c.3350A>G		p.K1117R		
TET2	ENST00000380013	NM_001127208.1	c.3365dupC		p.P1123TfsX7		
TET2	ENST00000380013	NM_001127208.1	c.3404G>A		p.C1135Y	YES	YES
TET2	ENST00000380013	NM_001127208.1	c.3493G>A		p.E1165K	YES	
TET2	ENST00000380013	NM_001127208.1	c.3499A>G		p.R1167G		
TET2	ENST00000380013	NM_001127208.1	c.3526A>G		p.R1176G	YES	
TET2	ENST00000380013	NM_001127208.1	c.3532delG		p.E1178KfsX48		
TET2	ENST00000380013	NM_001127208.1	c.3577T>A		p.C1193S	YES	
TET2	ENST00000380013	NM_001127208.1	c.3579T>G		p.C1193W		
TET2	ENST00000380013	NM_001127208.1	c.3589_3594+23delAAGTGGGTAAGTGTGACTTGATAAGCCT		Splice Disruption		
TET2	ENST00000380013	NM_001127208.1	c.3594+5G>A		Splice Disruption	YES	
TET2	ENST00000380013	NM_001127208.1	c.3646C>T		p.R1216*		
TET2	ENST00000380013	NM_001127208.1	c.3732_3733delCT		p.Y1245LfsX22	YES	
TET2	ENST00000380013	NM_001127208.1	c.3726_3757dupCAAACCTACTCGGAGCTTACCGAGACGCTGA		p.R1253TfsX11		
TET2	ENST00000380013	NM_001127208.1	c.3781C>T		p.R1261C	YES	
TET2	ENST00000380013	NM_001127208.1	c.3782G>A		p.R1261H	YES	
TET2	ENST00000380013	NM_001127208.1	c.3797A>G		p.N1266S		
TET2	ENST00000380013	NM_001127208.1	c.3803+5G>A		Splice Disruption	YES	
TET2	ENST00000380013	NM_001127208.1	c.3845G>A		p.G1282D		
TET2	ENST00000380013	NM_001127208.1	c.3869G>G		p.S1290*		
TET2	ENST00000380013	NM_001127208.1	c.3893G>A		p.C1298Y		
TET2	ENST00000380013	NM_001127208.1	c.3894dupT		p.K1299*		YES
TET2	ENST00000380013	NM_001127208.1	c.3899_3909delTTGCCAGAAGC		p.F1300*		
TET2	ENST00000380013	NM_001127208.1	c.3899T>C		p.F1300S		

TET2	ENST00000380013	NM_001127208.1	c.3941A>G	p.D1314G		
TET2	ENST00000380013	NM_001127208.1	c.3936delT	p.D1314MfsX49		
TET2	ENST00000380013	NM_001127208.1	c.3954+1G>T	Splice Disruption		
TET2	ENST00000380013	NM_001127208.1	c.3955-2A>G	Splice Disruption	YES	
TET2	ENST00000380013	NM_001127208.1	c.3965T>C	p.L1322P	YES	
TET2	ENST00000380013	NM_001127208.1	c.3986T>A	p.L1329Q	YES	
TET2	ENST00000380013	NM_001127208.1	c.3998T>A	p.M1333K		
TET2	ENST00000380013	NM_001127208.1	c.4015A>T	p.K1339*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4044+1G>T	Splice Disruption		
TET2	ENST00000380013	NM_001127208.1	c.4104C>G	p.F1368L		
TET2	ENST00000380013	NM_001127208.1	c.4138C>T	p.H1380Y		
TET2	ENST00000380013	NM_001127208.1	c.4151A>G	p.D1384G		
TET2	ENST00000380013	NM_001127208.1	c.4178C>T	p.T1393I	YES	
TET2	ENST00000380013	NM_001127208.1	c.4206_4212delCAATCGA	p.D1402EfsX44		
TET2	ENST00000380013	NM_001127208.1	c.4210C>T	p.R1404*	YES	YES
TET2	ENST00000380013	NM_001127208.1	c.4350delT	p.R1451GfsX7	YES	
TET2	ENST00000380013	NM_001127208.1	c.4393C>T	p.R1465*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4537+1G>A	Splice Disruption		
TET2	ENST00000380013	NM_001127208.1	c.4546C>T	p.R1516*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4579C>T	p.Q1527*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4609C>T	p.Q1537*		
TET2	ENST00000380013	NM_001127208.1	c.4639C>T	p.Q1547*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4748C>A	p.S1583*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4767T>A	p.Y1589*	YES	
TET2	ENST00000380013	NM_001127208.1	c.4740_4780dupAAACTCTTACACACTTCAGATACTATGGAAAGCACCAGCC	p.P1594QfsX16		
TET2	ENST00000380013	NM_001127208.1	c.4870C>T	p.Q1624*		
TET2	ENST00000380013	NM_001127208.1	c.4938_4948delGGGTCTCTATTinsCAG	p.G1647SfsX11	YES	
TET2	ENST00000380013	NM_001127208.1	c.5059C>T	p.Q1687*		
TET2	ENST00000380013	NM_001127208.1	c.5173C>A	p.P1725T		
TET2	ENST00000380013	NM_001127208.1	c.5197_5244delTTCATGGGAGCCACCTCTAGATTACAC	p.F1733PfsX3	YES	
TET2	ENST00000380013	NM_001127208.1	c.5303C>T	p.A1768V	YES	
TET2	ENST00000380013	NM_001127208.1	c.5363_5367dupACATG	p.L1790TfsX32		
TET2	ENST00000380013	NM_001127208.1	c.5393C>G	p.S1798*	YES	
TET2	ENST00000380013	NM_001127208.1	c.5518_5519delGCinsTG	p.A1840S		
TET2	ENST00000380013	NM_001127208.1	c.5582G>T	p.G1861V		
TET2	ENST00000380013	NM_001127208.1	c.5643T>A	p.H1881Q		
TET2	ENST00000380013	NM_001127208.1	c.5650A>G	p.T1884A	YES	YES
TET2	ENST00000380013	NM_001127208.1	c.5690T>G	p.R1897S	YES	
TET2	ENST00000380013	NM_001127208.1	c.5734C>G	p.H1912D	YES	
TET2	ENST00000380013	NM_001127208.1	c.5777G>A	p.R1926H		
TET2	ENST00000380013	NM_001127208.1	c.5885C>T	p.P1962L		
TP53	ENST00000445888	NM_000546.4	c.97T>C	p.S33P	YES	
TP53	ENST00000445888	NM_000546.4	c.376-1G>A	Splice Disruption	YES	
TP53	ENST00000445888	NM_000546.4	c.389T>A	p.L130H	YES	
TP53	ENST00000445888	NM_000546.4	c.395A>T	p.K132M		
TP53	ENST00000445888	NM_000546.4	c.406C>T	p.Q136*	YES	
TP53	ENST00000445888	NM_000546.4	c.437G>A	p.W146*	YES	
TP53	ENST00000445888	NM_000546.4	c.467G>C	p.R156P	YES	
TP53	ENST00000445888	NM_000546.4	c.473_483delGCGCCATGGCC	p.R158HfsX19	YES	
TP53	ENST00000445888	NM_000546.4	c.475G>C	p.A159P	YES	
TP53	ENST00000445888	NM_000546.4	c.476C>T	p.A159V		
TP53	ENST00000445888	NM_000546.4	c.487T>A	p.Y163H		
TP53	ENST00000445888	NM_000546.4	c.488A>G	p.Y163C		
TP53	ENST00000445888	NM_000546.4	c.493C>T	p.Q165*		
TP53	ENST00000445888	NM_000546.4	c.535C>T	p.H179Y	YES	YES
TP53	ENST00000445888	NM_000546.4	c.536A>G	p.H179R		
TP53	ENST00000445888	NM_000546.4	c.537T>G	p.H179Q	YES	
TP53	ENST00000445888	NM_000546.4	c.559+2T>G	Splice Disruption		
TP53	ENST00000445888	NM_000546.4	c.578A>G	p.H193R	YES	
TP53	ENST00000445888	NM_000546.4	c.659A>G	p.Y220C	YES	
TP53	ENST00000445888	NM_000546.4	c.660T>G	p.Y220*	YES	
TP53	ENST00000445888	NM_000546.4	c.672+1G>A	Splice Disruption		
TP53	ENST00000445888	NM_000546.4	c.713G>A	p.C238Y		
TP53	ENST00000445888	NM_000546.4	c.734G>A	p.G245D	YES	
TP53	ENST00000445888	NM_000546.4	c.734G>C	p.G245A	YES	
TP53	ENST00000445888	NM_000546.4	c.742C>T	p.R248W	YES	YES
TP53	ENST00000445888	NM_000546.4	c.743G>A	p.R248Q		
TP53	ENST00000445888	NM_000546.4	c.783dupT	p.G262WfsX2		
TP53	ENST00000445888	NM_000546.4	c.814G>A	p.V722M	YES	YES
TP53	ENST00000445888	NM_000546.4	c.862_863delAAinsT	p.N288HfsX57	YES	
TP53	ENST00000445888	NM_000546.4	c.917_919+13delGAGGTAAGCAAGCAGGinsAGT	Splice Disruption		
TP53	ENST00000445888	NM_000546.4	c.920-1G>A	Splice Disruption		
TP53	ENST00000445888	NM_000546.4	c.920-2A>G	Splice Disruption	YES	
TP53	ENST00000445888	NM_000546.4	c.949C>T	p.Q317*		
TP53	ENST00000445888	NM_000546.4	c.1007delA	p.E336GfsX9		
KDM6A	ENST00000377967	NM_021140.2	c.1751C>T	p.T584M		
KDM6A	ENST00000377967	NM_021140.2	c.2331T>A	p.N777K		
KDM6A	ENST00000377967	NM_021140.2	c.4093A>G	p.T1365A		

Supplementary Table 5: List of mutations proven to be or previously reported as germline and not included in dbSNP.

Gene Name	Ensembl Reference Transcript	RefSeq Reference Transcript	DNA Mutation	Protein Mutation	Reason for Exclusion	Present in Two or More Samples
ASXL1	ENST00000375687	NM_015338.5	c.1181G>A	p.R394H	Identified in Germline	
ASXL1	ENST00000375687	NM_015338.5	c.1465C>G	p.R489G	Identified in Germline	
ASXL1	ENST00000375687	NM_015338.5	c.1720-3C>G	Splice Disruption	Identified in Germline	
ASXL1	ENST00000375687	NM_015338.5	c.2468T>C	p.L823S	Identified in Germline	
ASXL1	ENST00000375687	NM_015338.5	c.2957A>G	p.N986S	Identified in Germline	YES
ASXL1	ENST00000375687	NM_015338.5	c.3498C>G	p.S1166R	Identified in Germline	YES
ASXL1	ENST00000375687	NM_015338.5	c.3745A>G	p.M1249V	Previously Reported as Germline	
ASXL1	ENST00000375687	NM_015338.5	c.3935C>T	p.A1312V	Identified in Germline	
ASXL1	ENST00000375687	NM_015338.5	c.4099G>A	p.V1367I	Identified in Germline	
ASXL1	ENST00000375687	NM_015338.5	c.4189G>A	p.G1397S	Identified in Germline	
CDH1	ENST00000261769	NM_004360.3	c.1849G>A	p.A617T	Identified in Germline	YES
CDKN2A	ENST00000304494	NM_000077.3	c.47T>C	p.L16P	Identified in Germline	
CDKN2A	ENST00000304494	NM_000077.3	c.146T>C	p.I49T	Identified in Germline	
CDKN2A - p16INK4A / p14ARF	ENST00000304494 / ENST00000361570	NM_000077.3 / NM_058195.2	c.205G>A / c.371G>A	p.E69K / p.G124E	Identified in Germline	
CDKN2A	ENST00000361570	NM_058195.2	c.290G>A	p.G97E	Identified in Germline	
EGFR	ENST00000275493	NM_005228.3	c.2369C>T	p.T790M	Identified in Germline	
ETV6	ENST00000266427	NM_001987.4	c.672C>G	p.H224Q	Identified in Germline	
EZH2	ENST00000320356	NM_004456.3	c.965A>G	p.N322S	Identified in Germline	
MET	ENST00000318493	NM_001127500.1	c.504G>T	p.E168D	Identified in Germline	YES
PTEN	ENST00000371953	NM_000314.4	c.235G>A	p.A79T	Identified in Germline	
RUNX1	ENST00000300305	NM_001754.4	c.155T>A	p.M52K	Identified in Germline	
RUNX1	ENST00000300305	NM_001754.4	c.167T>C	p.L56S	Previously Reported as Germline	YES
RUNX1	ENST00000300305	NM_001754.4	c.733C>T	p.P245S	Identified in Germline	
TET2	ENST00000380013	NM_001127208.1	c.100C>T	p.L34F	Previously Reported as Germline	YES
TET2	ENST00000380013	NM_001127208.1	c.5666C>T	p.P1889L	Identified in Germline	
TET2	ENST00000380013	NM_001127208.1	c.1285G>A	p.G429R	Identified in Germline	
TET2	ENST00000380013	NM_001127208.1	c.1379C>T	p.S460F	Identified in Germline	
TET2	ENST00000380013	NM_001127208.1	c.2599T>C	p.Y867H	Previously Reported as Germline	YES
TET2	ENST00000380013	NM_001127208.1	c.3251A>C	p.Q1084P	Previously Reported as Germline	YES
TET2	ENST00000380013	NM_001127208.1	c.3255_3257delAAC	p.T1085del	Identified in Germline	
TET2	ENST00000380013	NM_001127208.1	c.5152G>T	p.V1718L	Identified in Germline	YES
TET2	ENST00000380013	NM_001127208.1	c.5167C>T	p.P1723S	Previously Reported as Germline	YES
TET2	ENST00000380013	NM_001127208.1	c.5449C>A	p.H1817N	Identified in Germline	
TP53	ENST00000445888	NM_000546.4	c.704A>G	p.N235S	Identified in Germline	
KDM6A	ENST00000377967	NM_021140.2	c.1843C>G	p.L615V	Identified in Germline	

Supplementary Table 6: Coincidence of Mutations: This table shows the number of samples that had mutations in any given pair of genes. The gray-shaded cells on the diagonal show the number of samples that only have mutations in that gene. Cells in the top row list the total number of samples with a mutation in the gene listed in that column. Samples with more than two mutations will be counted in more than one cell per row and column. Therefore, the total number of mutation pairs (listed in the bottom row) may be greater than number of mutated samples shown in the top row.

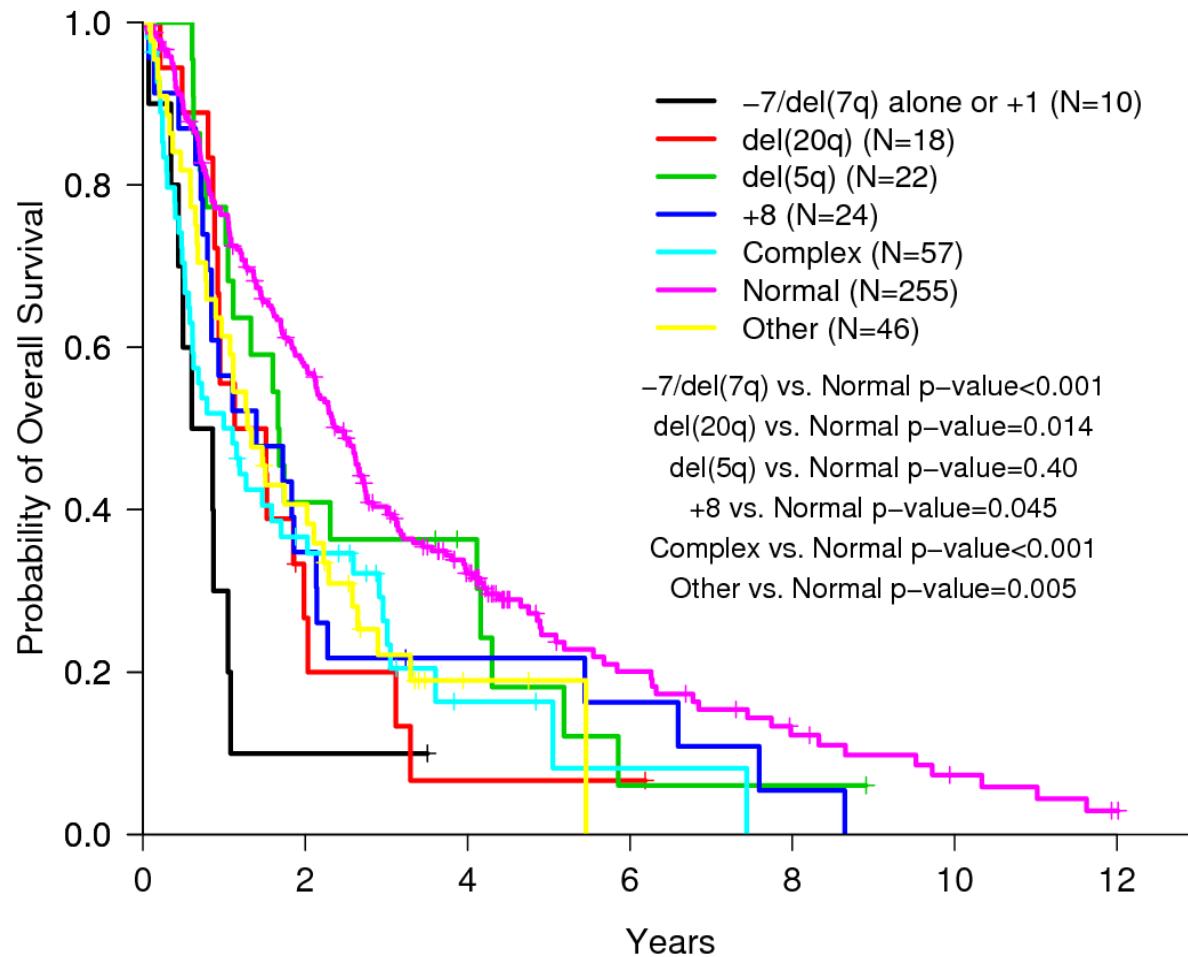
N	90	63	38	33	28	19	16	12	10	9	8	6	4	3	3	2	1	1
	TET2	ASXL1	RUNX1	TP53	EZH2	NRAS	JAK2	ETV6	CBL	IDH2	NPM1	IDH1	KRAS	GNAS	PTPN11	BRAF	PTEN	CDKN2A
TET2	46	18	12	7	7	7	3	2	7	1	1		2				1	
ASXL1	18	22	13	1	14	4	3	2	5	4		3	1					
RUNX1	12	13	14		8	6			1				1		1			
TP53	7	1		23	2			1									1	
EZH2	7	14	8	2	6	2		1	1	1			1					
NRAS	7	4	6		2	3						1	1		1			
JAK2	3	3					6	1		1								
ETV6	2	2		1	1		1	4						1	1			
CBL	7	5	1		1				2	2								
IDH2	1	4			1		1		2	2				1				
NPM1	1										6							
IDH1		3			1	1						3						
KRAS	2	1	1			1							2					
GNAS								1		1				1				
PTPN11			1			1		1							1			
BRAF																2		
PTEN	1			1														1
CDKN2A																		1
TOTAL	115	90	56	36	43	25	14	13	18	12	8	8	7	3	4	2	2	1

Supplementary Table 7: Association of Karyotype with Mutations and Overall Survival. Table A shows the distribution and frequency of karyotype abnormalities by mutation status. Figures B and C are Kaplan-Meier survival curves illustrating the overall survival of patients based on karyotype and karyotype group as defined in the IPSS at the time of sample collection.

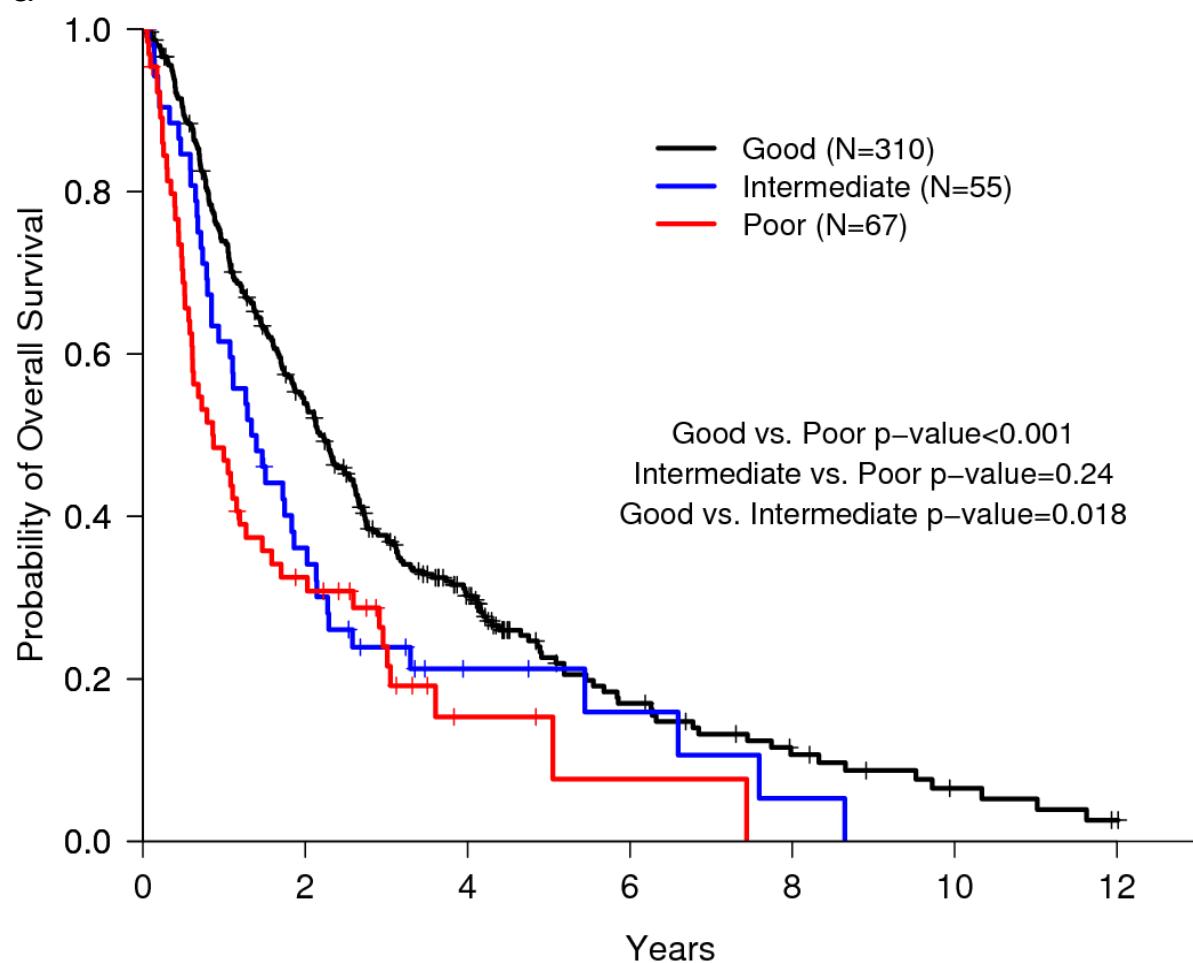
A.

Karyotype	-7/del(7q) (alone or +1)	del(20q)	del(5q)	+8	Complex	Normal	Other	p-value
N	10	18	22	24	57	255	46	
At least 1 Mutation	7 (70)	11 (61)	8 (36)	13 (54)	32 (56)	133 (52)	20 (43)	0.44
No Mutation	3 (30)	7 (39)	14 (64)	11 (46)	25 (44)	122 (48)	26 (57)	
TET2+	1 (10)	2 (11)	3 (14)	5 (21)	3 (5)	68 (27)	7 (15)	0.005
TP53+	0 (0)	0 (0)	1 (5)	0 (0)	26 (46)	5 (2)	1 (2)	<0.001
EZH2+	0 (0)	2 (11)	1 (5)	2 (8)	0 (0)	20 (8)	3 (7)	0.23

B.



c.

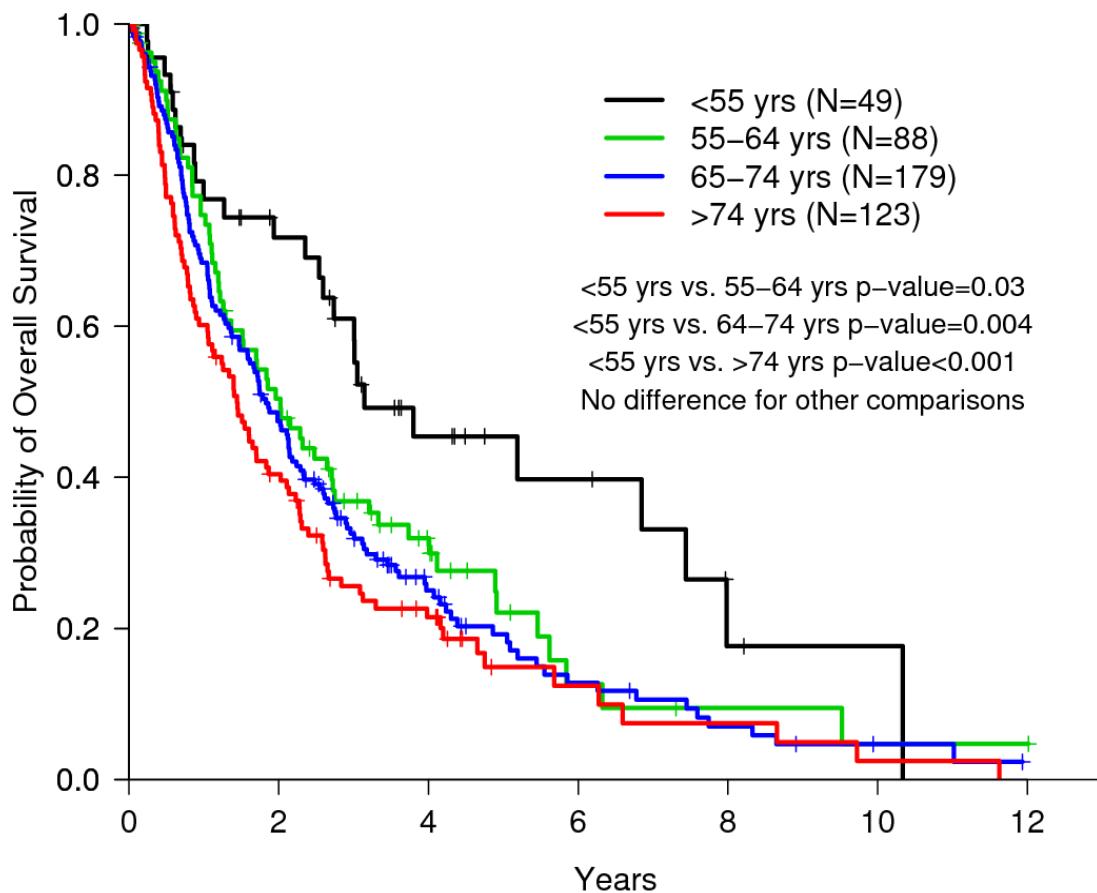


Supplementary Table 8: Association of Age with Mutations and Overall Survival. Table A shows the distribution and frequency of mutation and karyotype status by patient age. Figure B is a Kaplan-Meier survival curve illustrating the overall survival of patients based on age group at the time of sample collection. Pairwise comparisons demonstrate that the survival of patients aged less than 55 is significantly different from that in every other age group, but that the survival of patients in these older age groups is not significantly different from each other.

A.

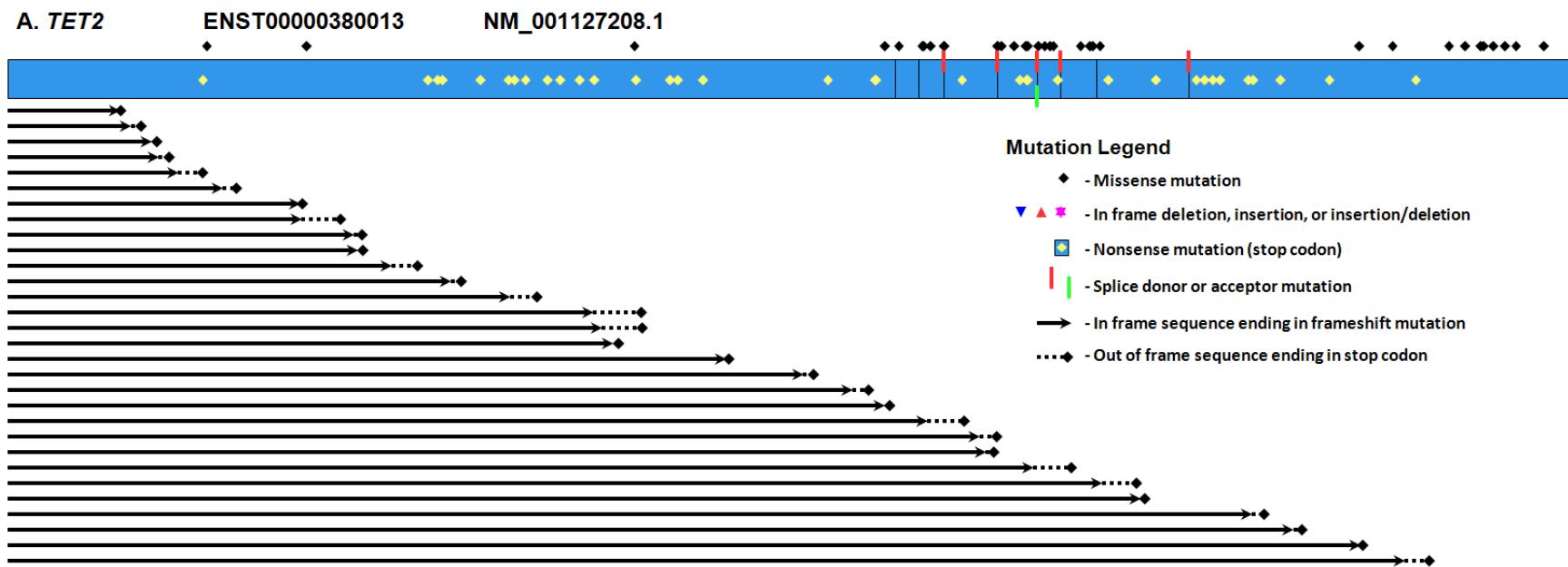
Age Group	N (%)				p-value
	<55 yrs.	55-64 yrs	65-74 yrs.	≥75 yrs.	
N	49	88	179	123	
At least 1 Mutation	16 (33)	45 (51)	94 (53)	71 (58)	0.02
No Mutation	33 (67)	43 (49)	85 (47)	52 (42)	
No Mut/Normal Karyo	17 (35)	28 (33)	48 (27)	29 (24)	0.07
Other	31 (65)	58 (67)	128 (73)	93 (76)	

B.

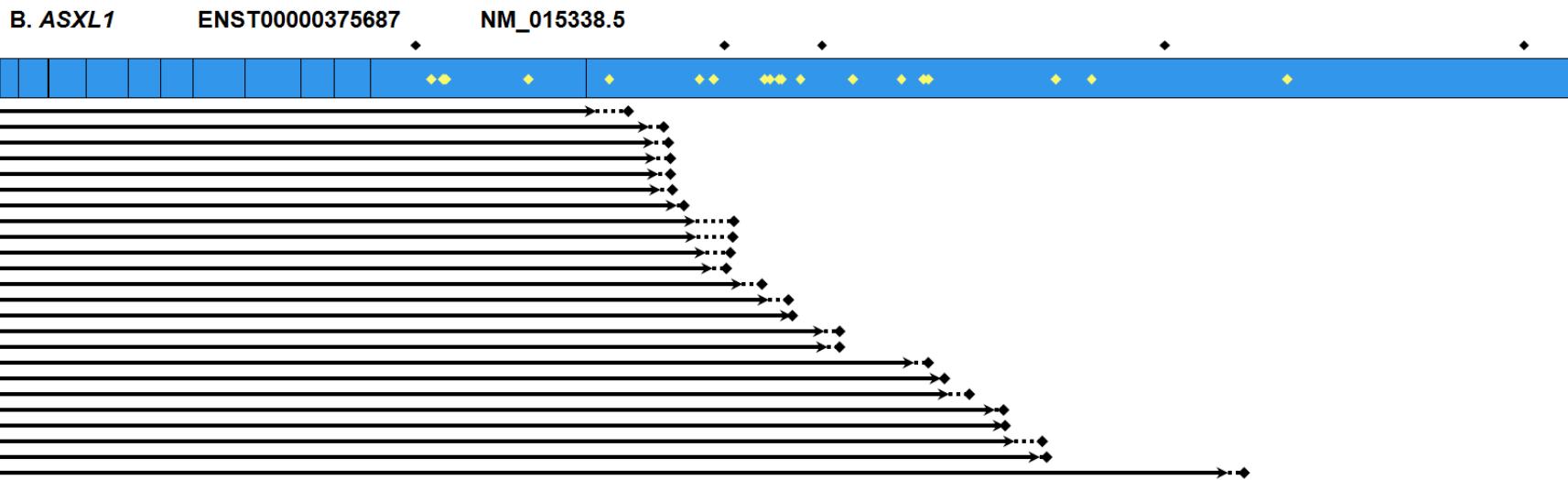


Supplementary Figures

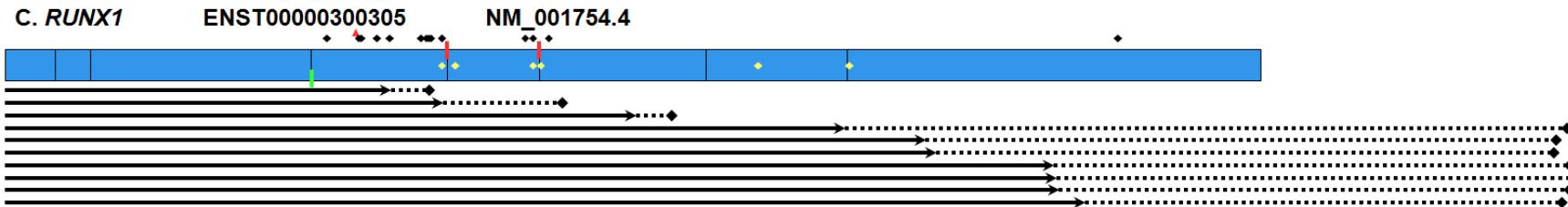
Supplementary Figure 1: Location and Type of Mutations within Gene Products. These figures show where amino acid altering mutations begin in each of six gene products. The gene product is shown as blue rectangles joined at exon boundaries. The mutation types are indicated by symbols described in the legend shown in Figure A.



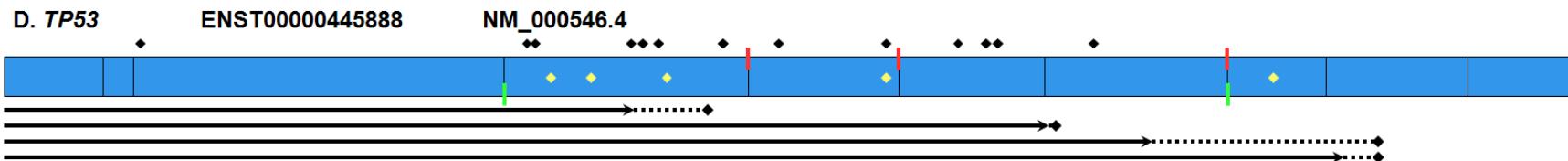
Stop codon and frameshift mutations predicted to truncate the *TET2* gene product occur throughout its length. Missense and splice site mutations are clustered in one of two highly conserved C-terminus domains associated with the catalytic activity of *TET2*. Nearly a quarter of patients had two or more distinct mutations in *TET2* suggesting biallelic abnormalities. Mitotic recombination leading to copy-neutral loss of heterozygosity has been seen recurrently in *TET2*-mutated MDS samples. Together, these patterns of mutation are consistent with a loss of function of this gene.



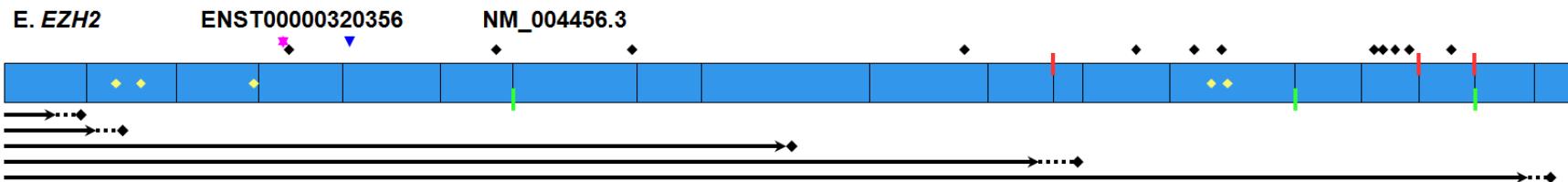
Mutations of *ASXL1* are almost exclusively stop codons and frameshifts predicted to prematurely truncate its protein product. The bulk of these mutations occur in the terminal exon and are therefore, unlikely to induce nonsense-mediated decay. Mutated *ASXL1* genes may produce abnormal proteins that retain their N-terminal domains which could lead to a gain-of-function or dominant-negative activity.



The *RUNX1* protein contains two functionally important regions, a more proximal RUNT DNA binding domain and a more distal protein interaction domain. In frame mutations of *RUNX1* are clustered in the RUNT domain. Similar mutations have been shown to produce proteins with dominant negative activity. Frameshift, nonsense, and splice site mutations appear to spare the proximal portion of the gene, although beyond a loss of normal activity, the functional impact of these mutations is unclear.



As with *RUNX1*, mutations of *TP53* mostly spared the proximal portion of the gene. Several patients had two distinct *TP53* mutations or a single mutation and loss of chromosome 17p where the *TP53* gene resides. Most of the missense mutations we identified occurred in the central DNA binding domain and have previously been associated with a loss of function.

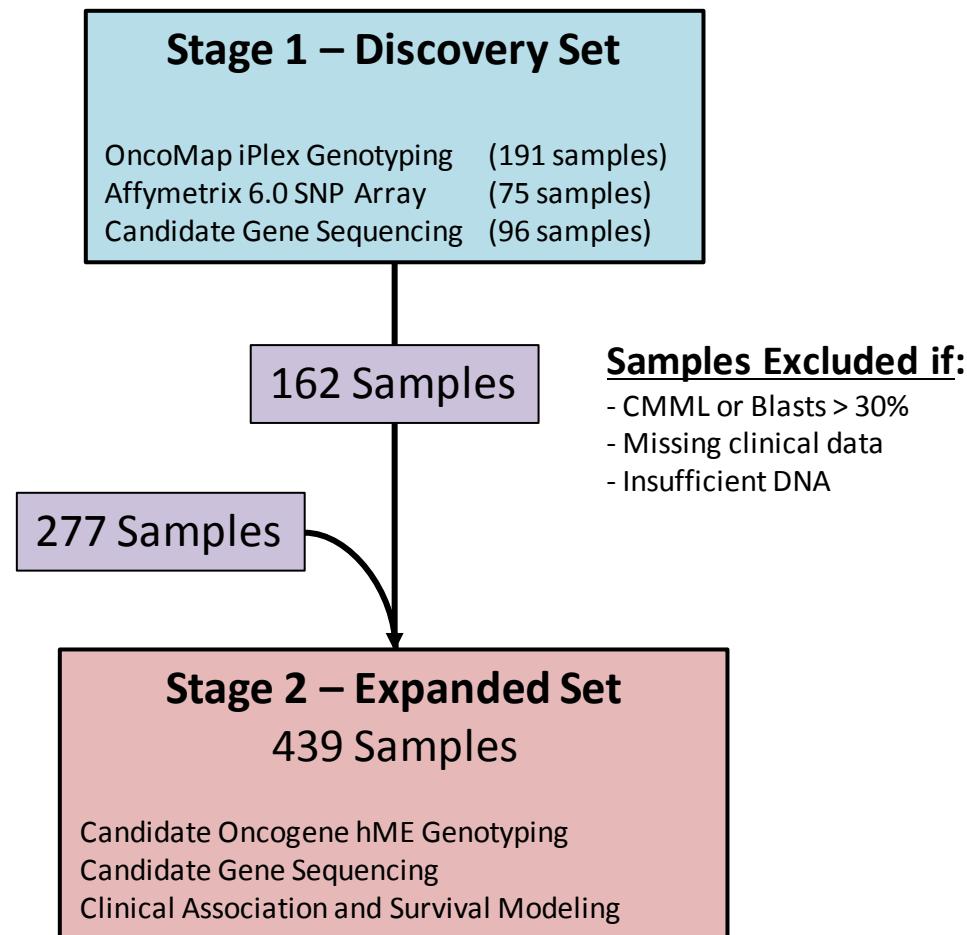


Mutations of *EZH2* spanned the length of the gene. Several missense mutations clustered in the C-terminal SET domain responsible for methyltransferase activity. This pattern is consistent with a loss of function as the selection driver for *EZH2* mutations.

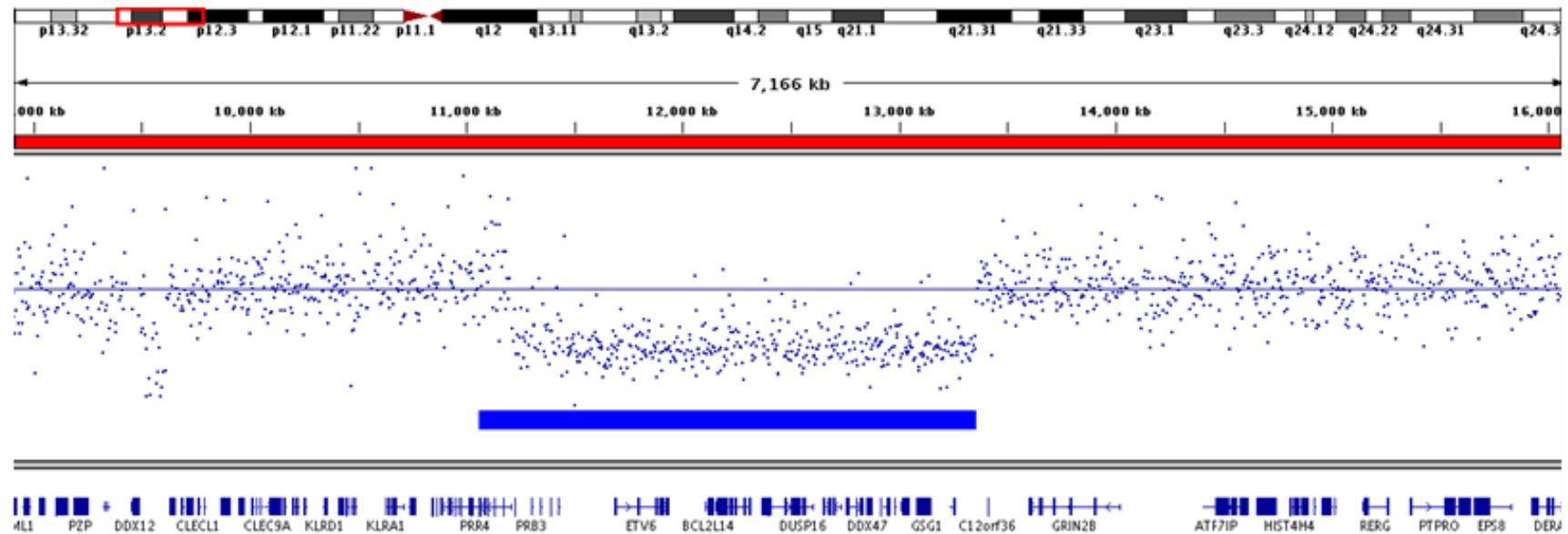


Most of the *ETV6* mutations we identified occur in the sterile alpha motif domain responsible for dimerization with other *ETV6* molecules or other ets proteins.

Supplementary Figure 2: Diagram of Two Stage Study Design.

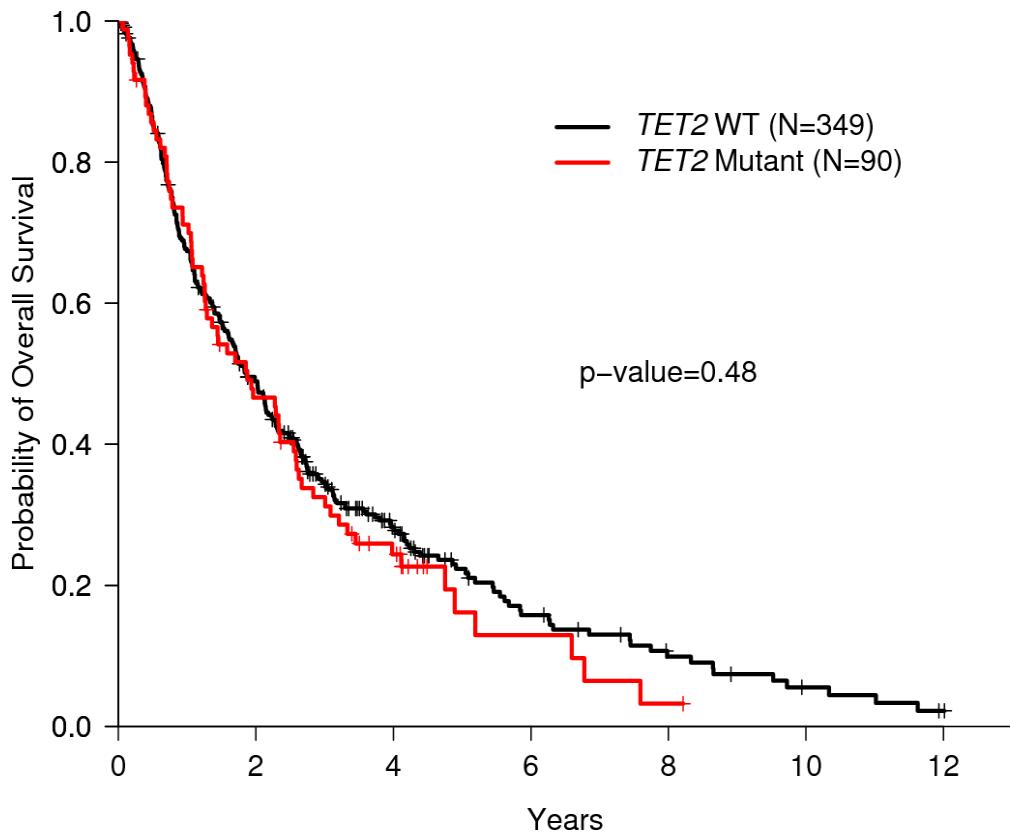


Supplementary Figure 3: Copy number results from an MDS bone marrow sample genotyped on an Affymetrix 6.0 SNP array demonstrates an interstitial 2.3 megabase deletion on chromosome 12p that includes the *ETV6* gene locus.

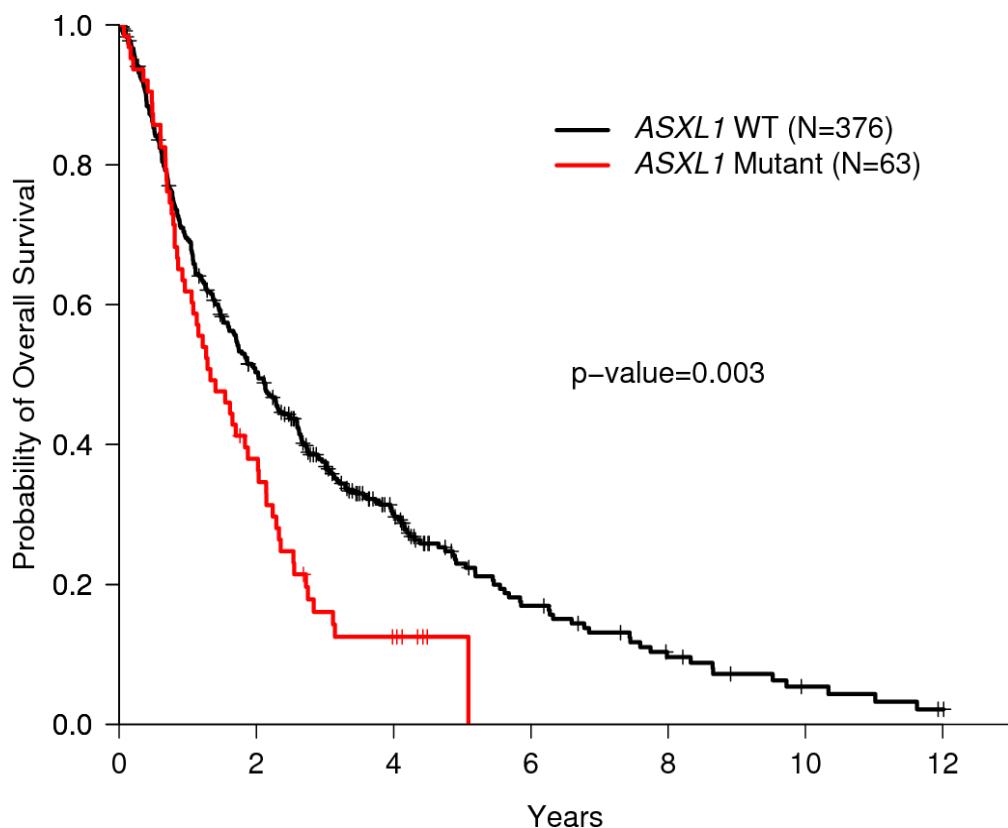


Supplementary Figure 4: Kaplan-Meier Survival Curves for Mutations. Each panel shows the proportion of surviving patients with a given mutation (red line) compared to patients without that mutation (black line). The numbers of unmutated vs. mutated cases are shown in parentheses, respectively. The log-rank *p*-value is provided comparing the survival of the two groups in each panel.

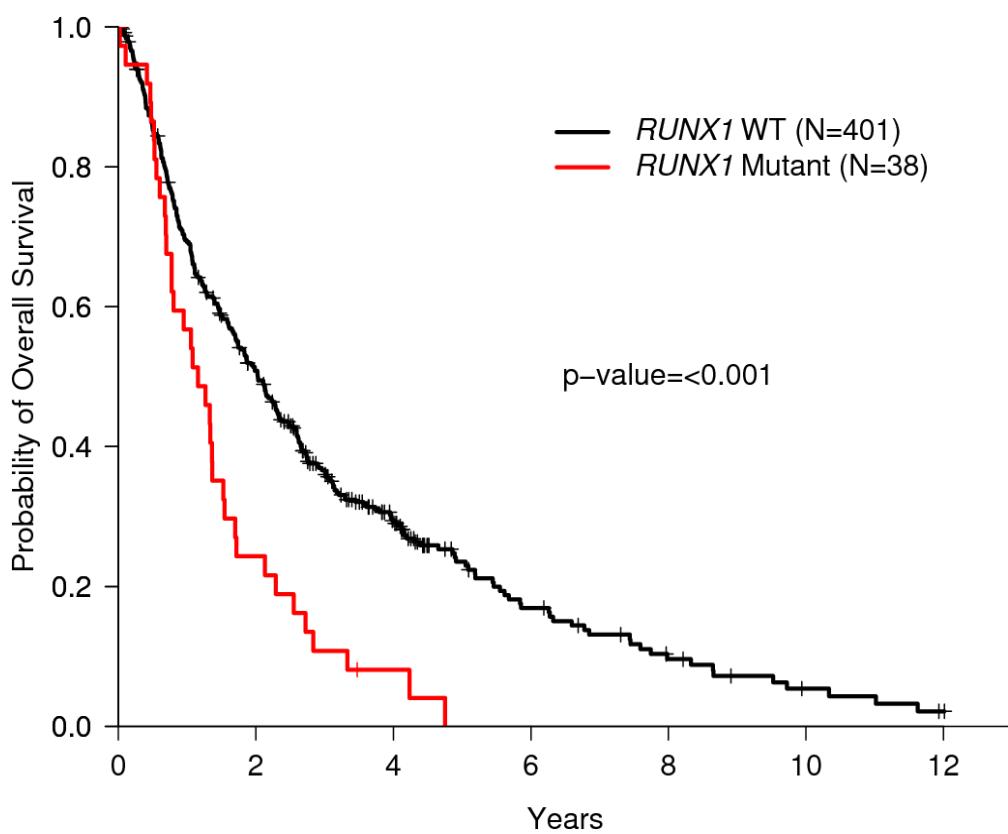
A.



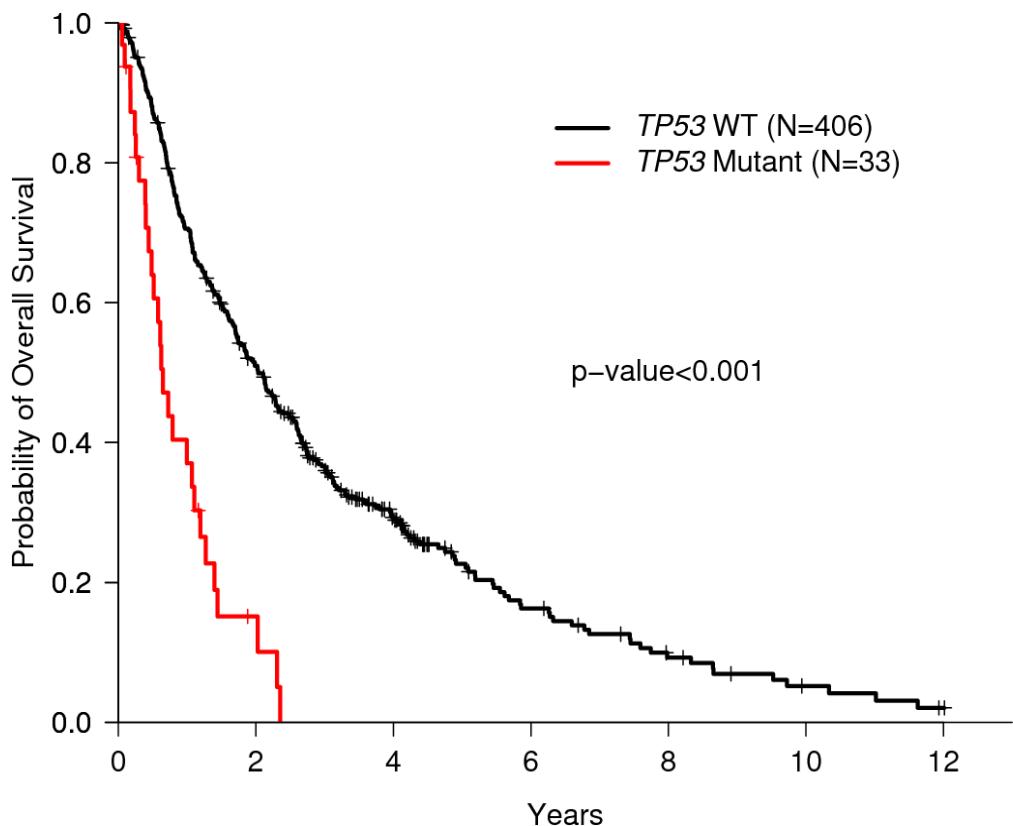
B.



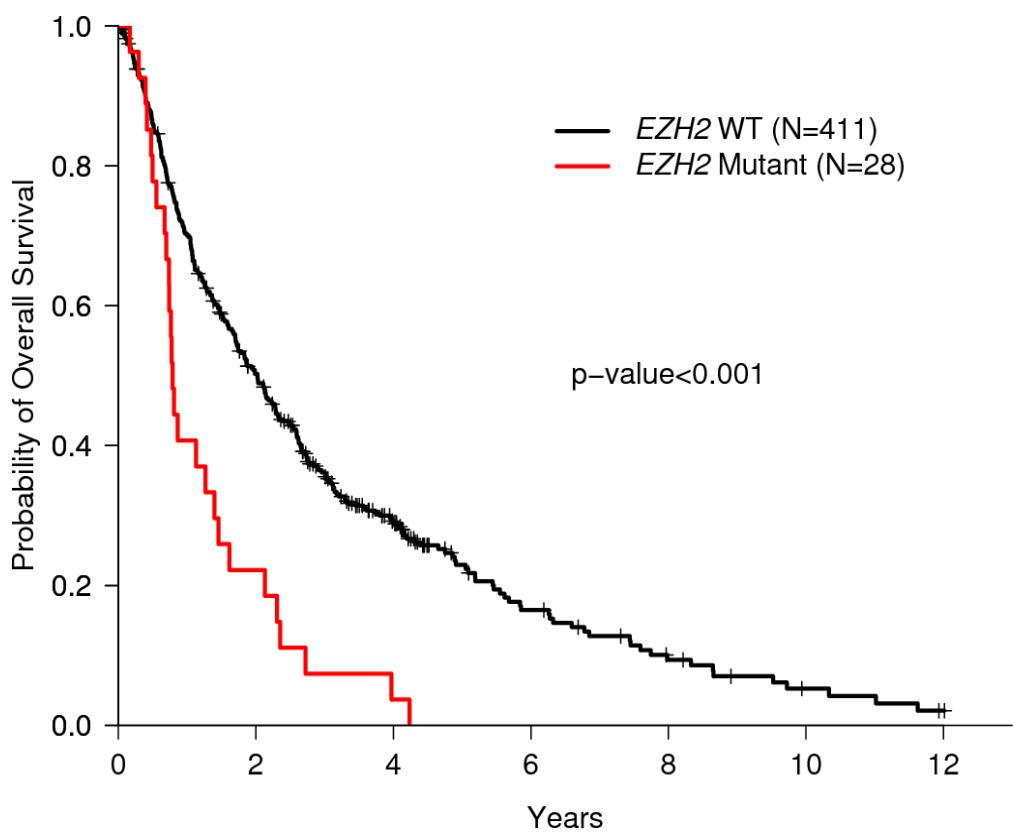
C.



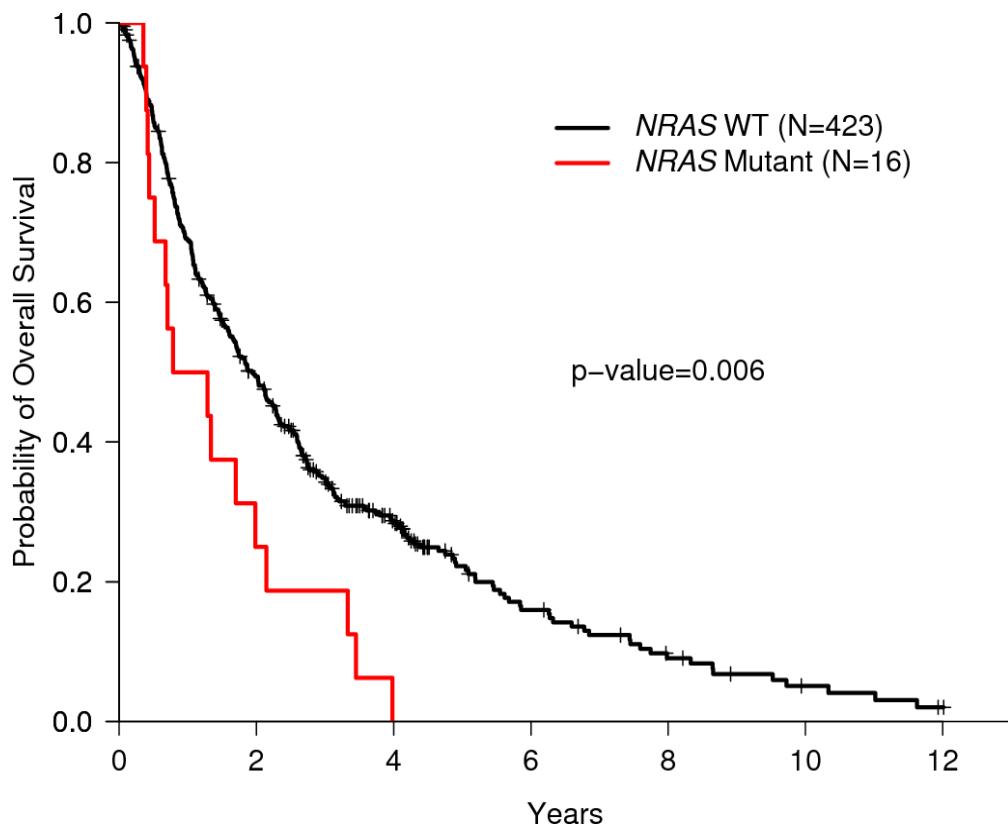
D.



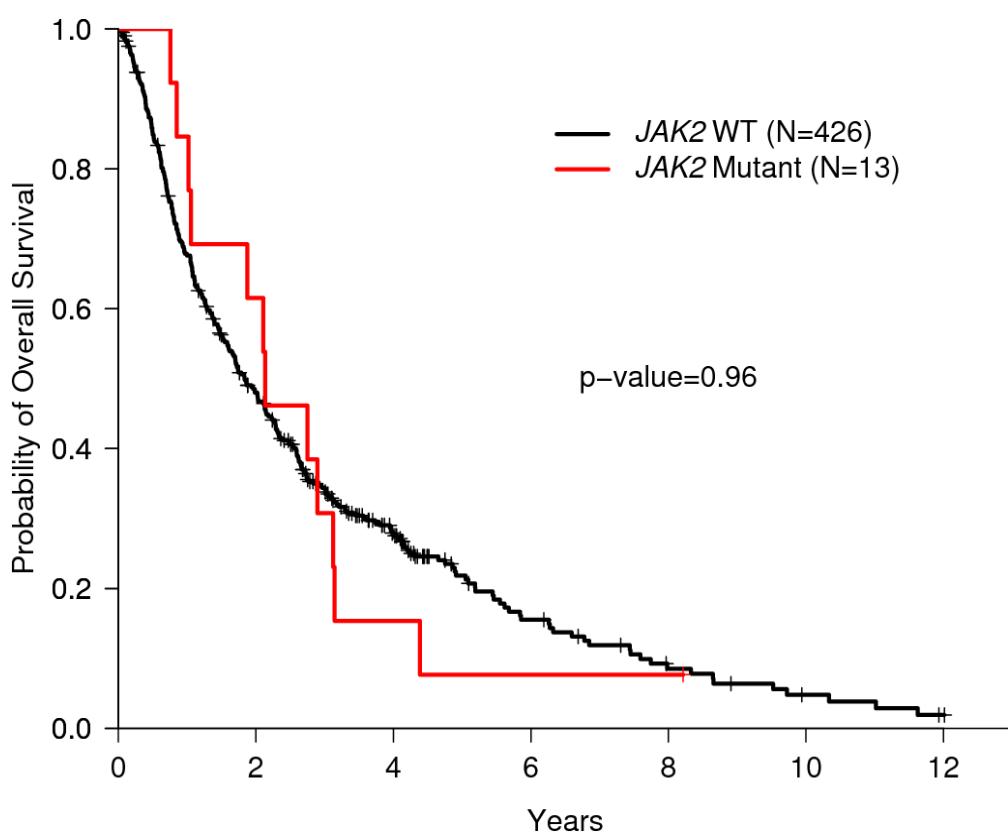
E.



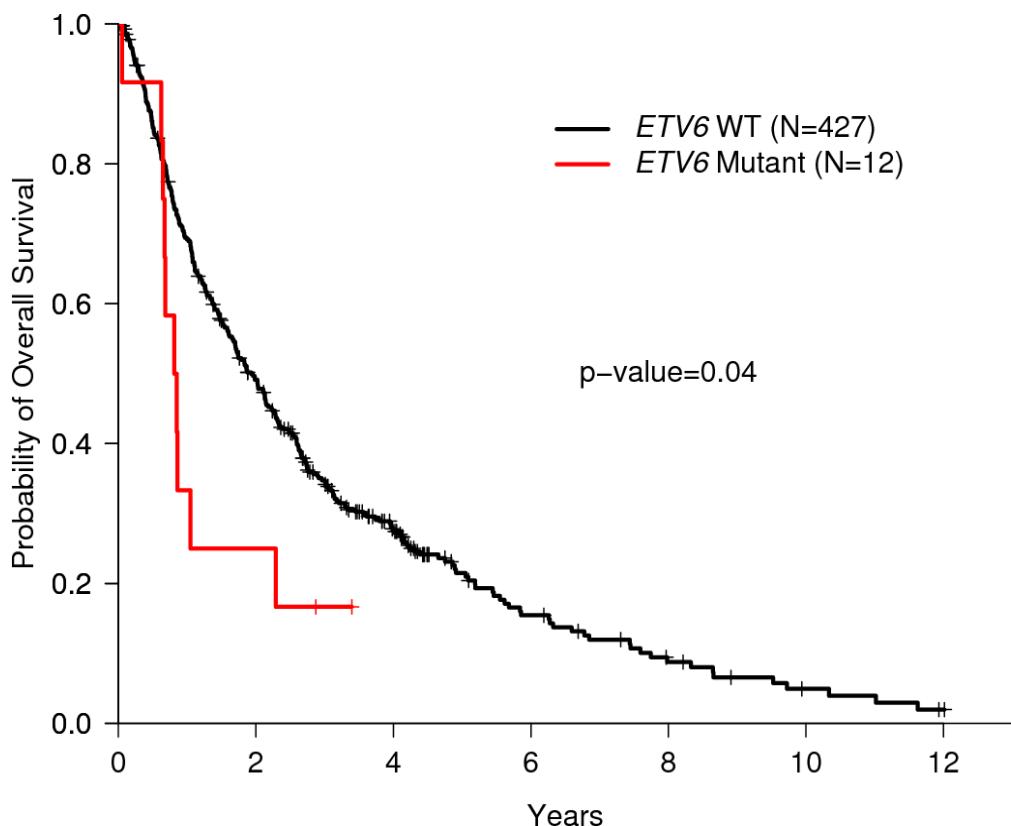
F.



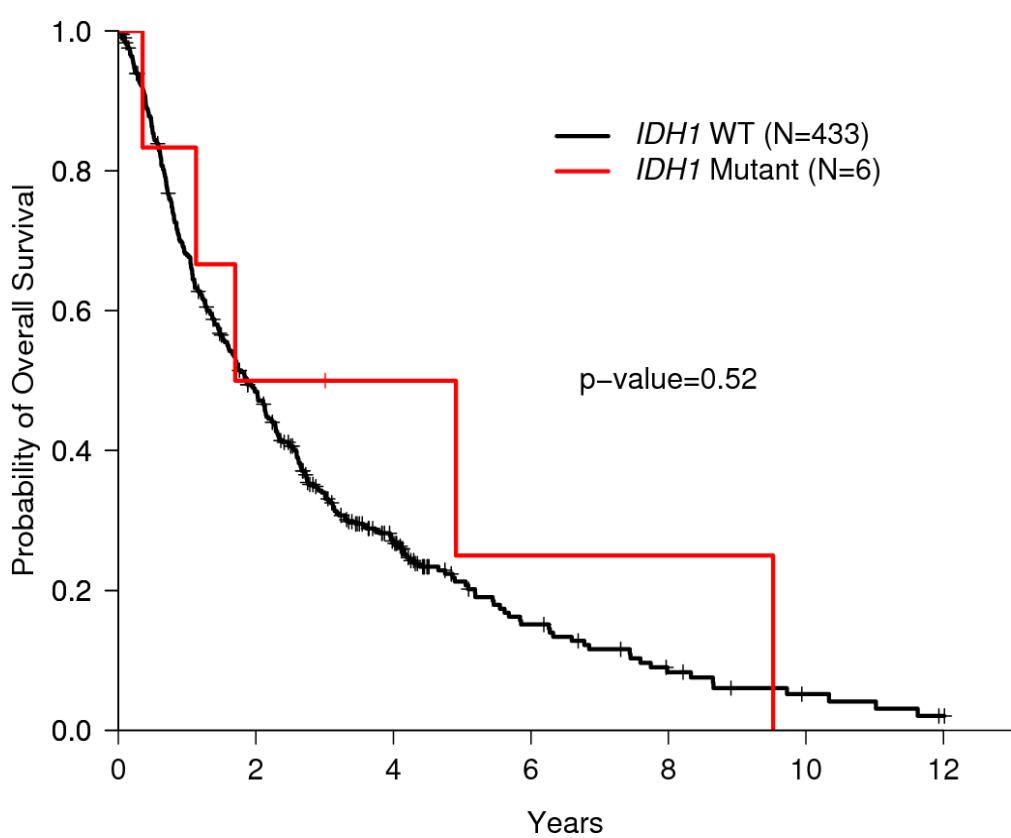
G.



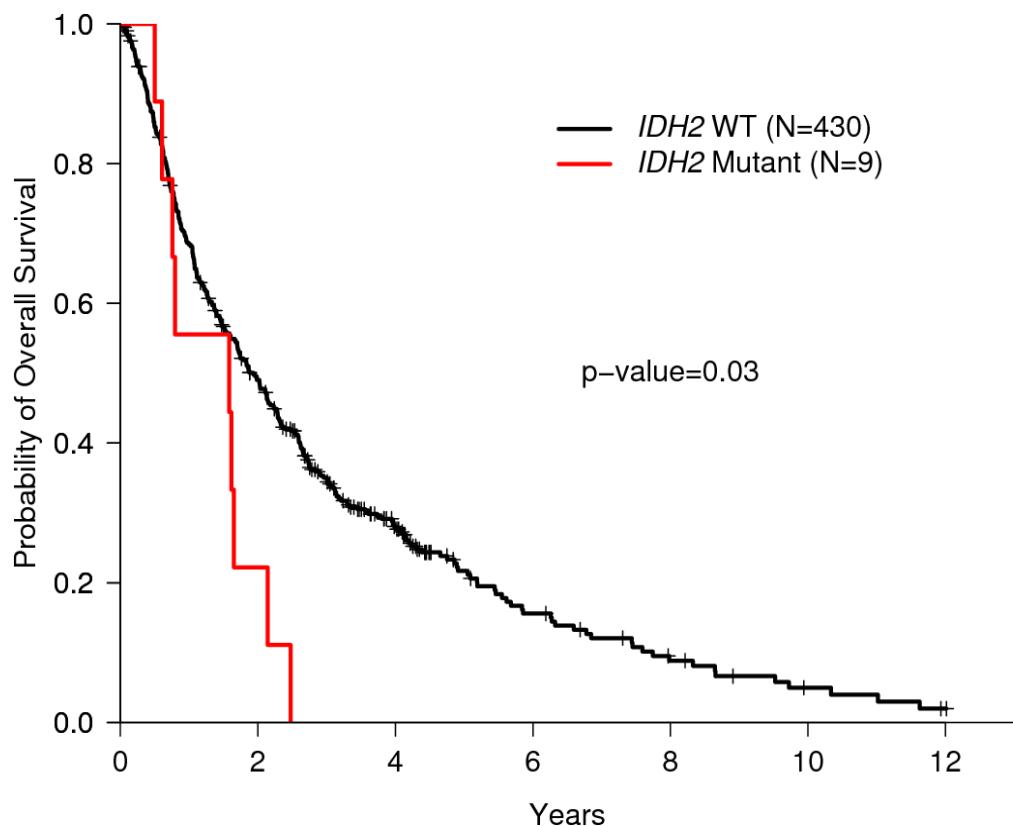
H.



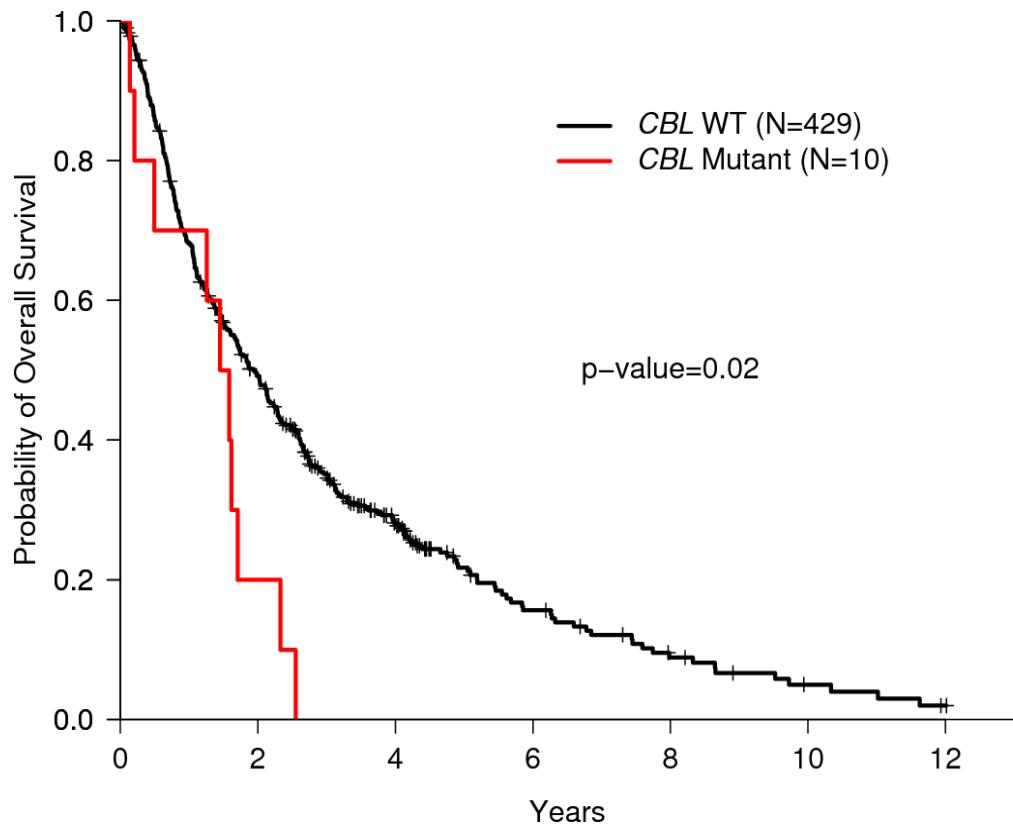
I.



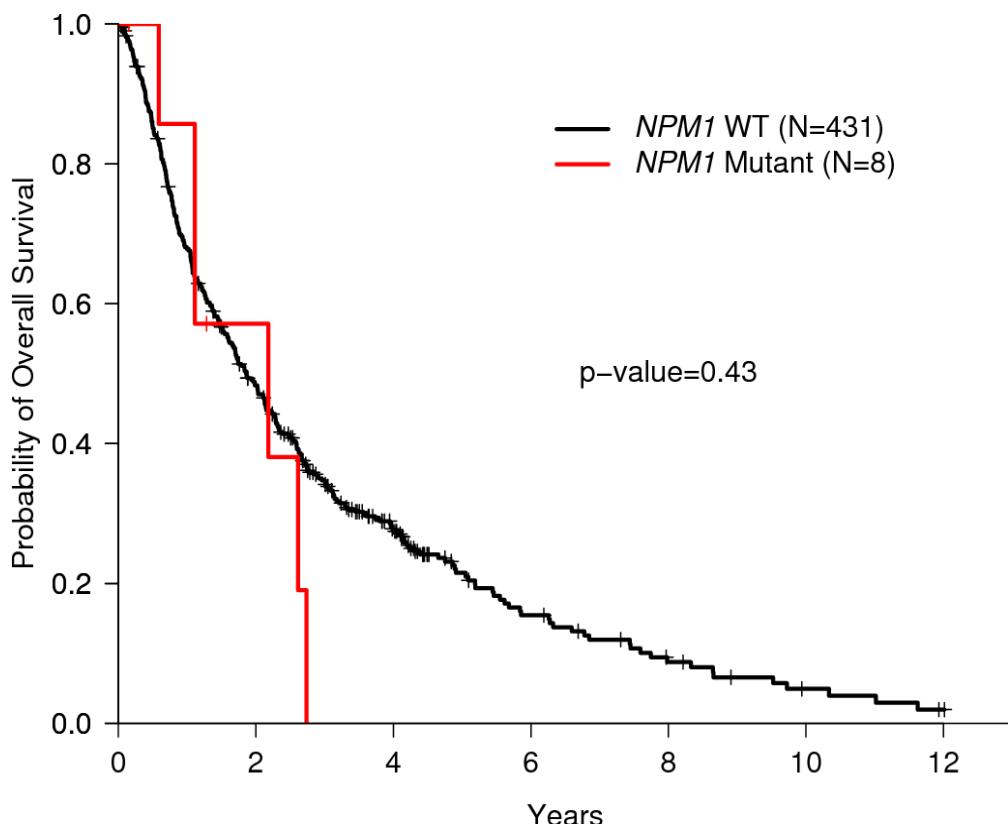
J.



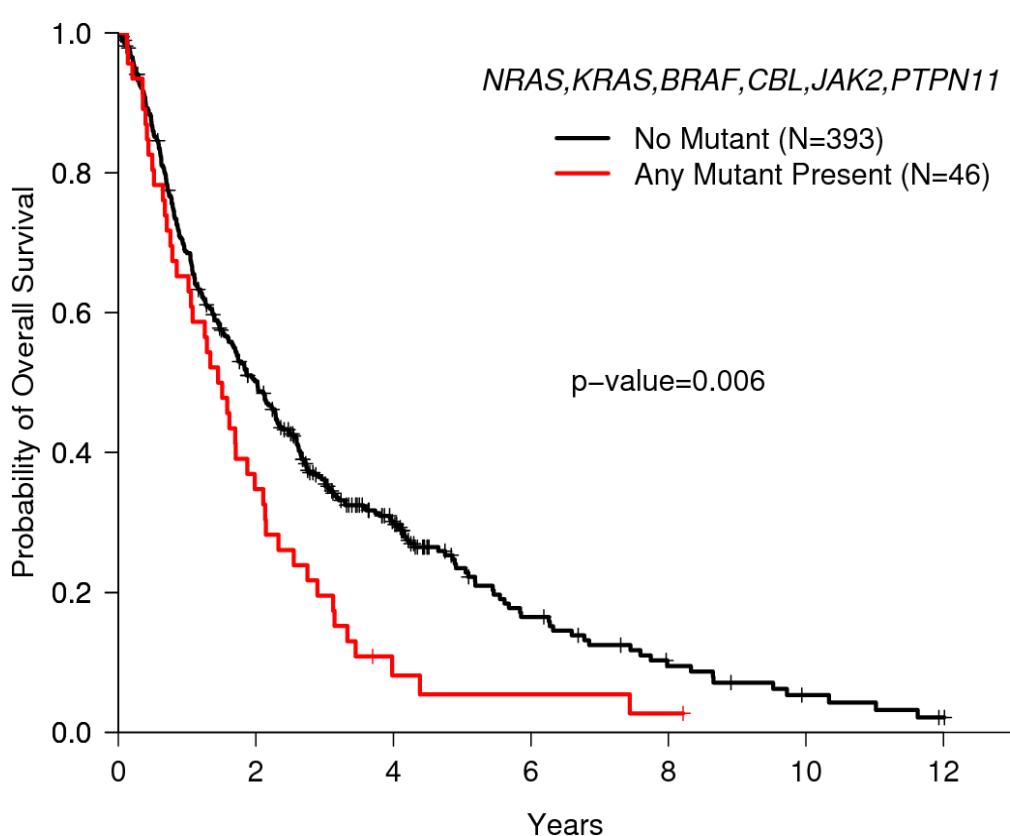
K.



L.

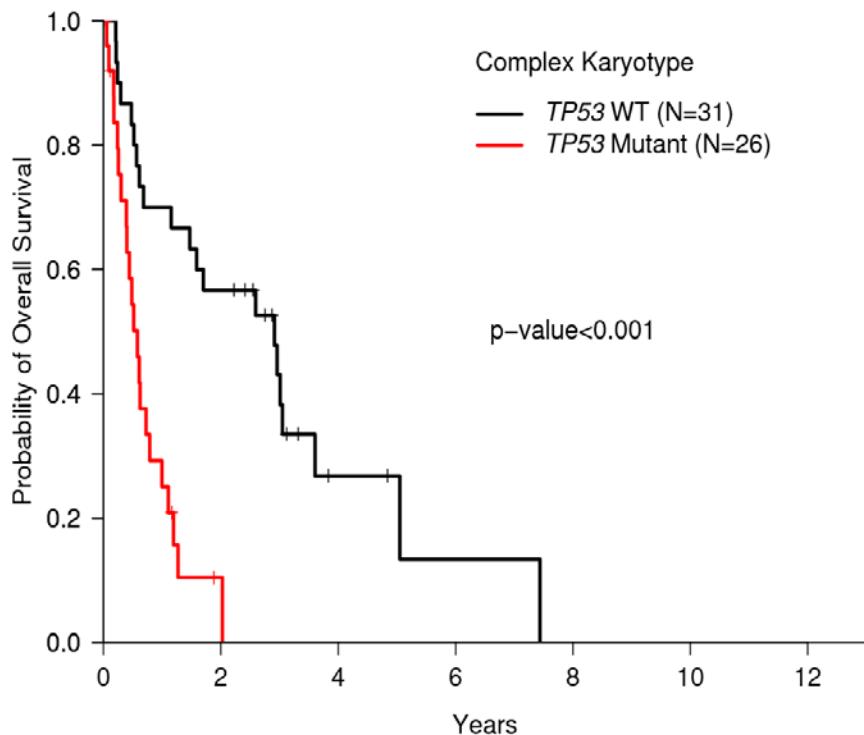


M.

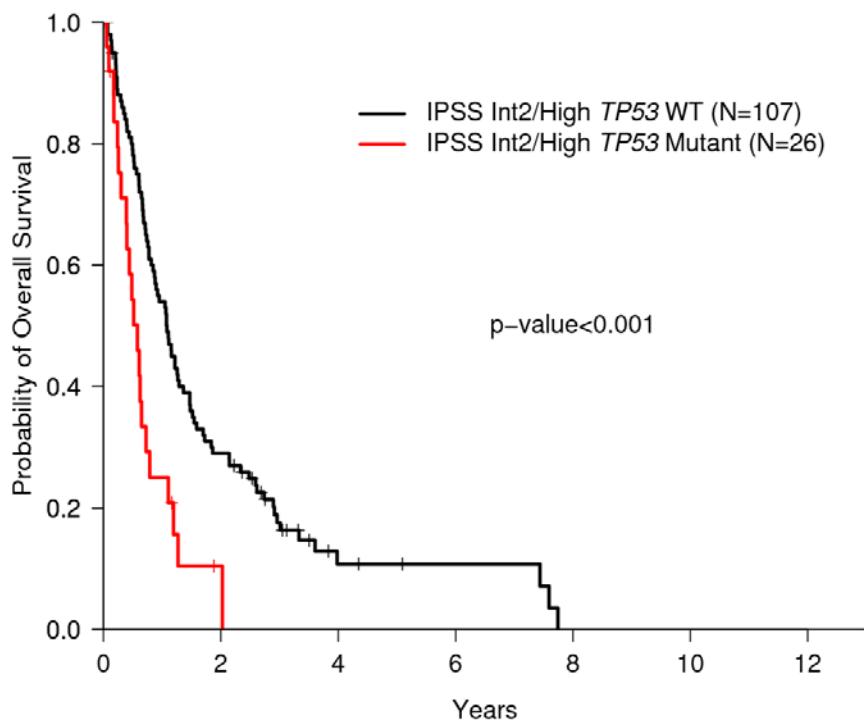


Supplementary Figure 5: Survival of Patients with Complex Cytogenetics or Higher Risk IPSS Groups Stratified by TP53 Mutation Status. A) Patients with complex cytogenetics and *TP53* mutation have significantly shorter overall survival than patients with complex cytogenetics and no *TP53* mutation. B) Patients with increased IPSS risk and *TP53* mutation have significantly shorter overall survival than patients with increased IPSS risk and no *TP53* mutation.

A.

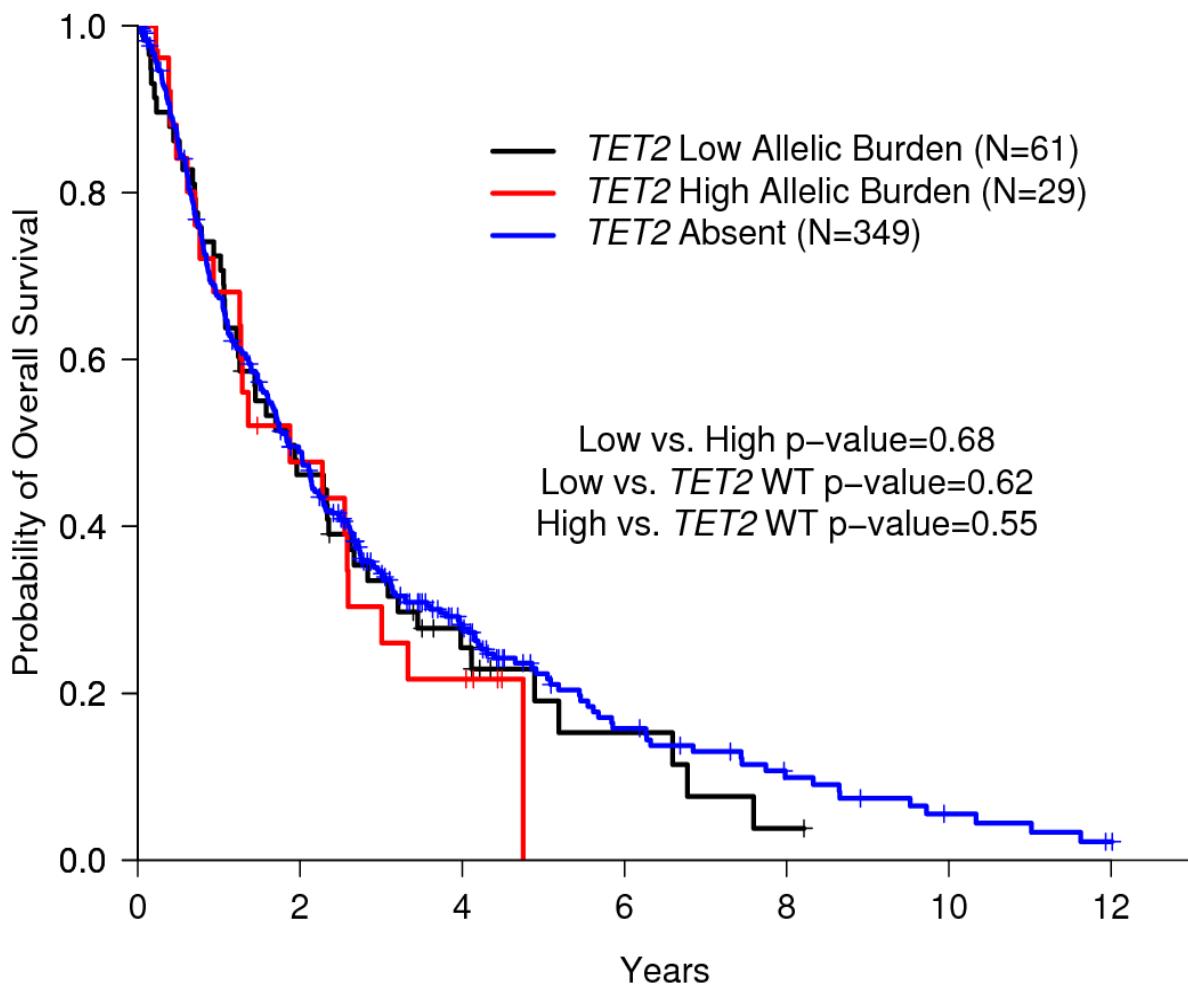


B.

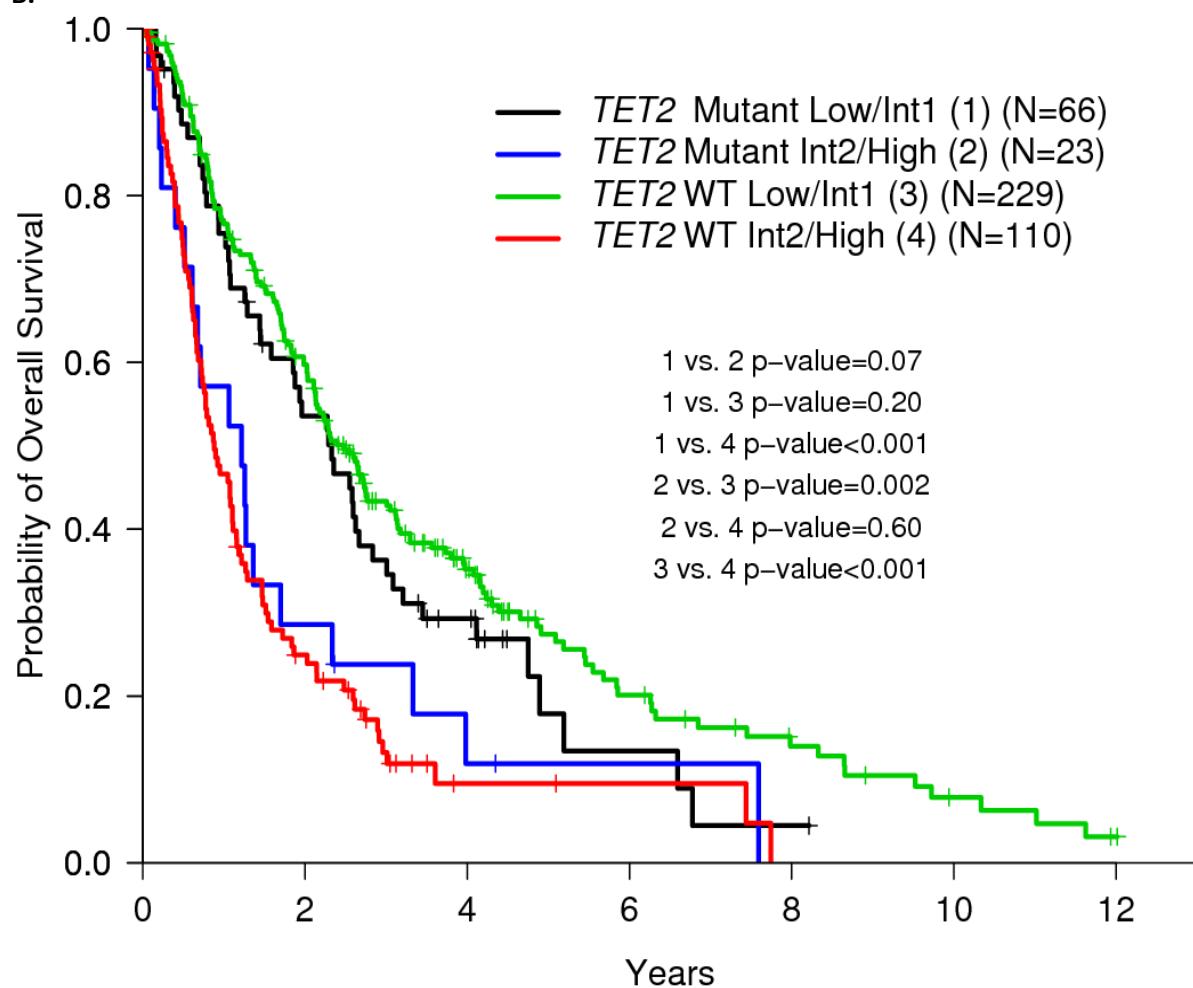


Supplementary Figure 6: Survival of Patients with *TET2* Mutations Stratified by Mutant Allele Burden and Prognostic Risk. In panel A, the overall survival of patients with more than two *TET2* mutations or a single *TET2* mutation present at twice the frequency of the wild type allele (red line) is compared to the survival of patients with a single *TET2* mutation and lower mutant allele burden (blue line) and to the survival of patients without a *TET2* mutation (black line). The proportion of mutant alleles in each sample was determined by dividing the area of the mutant allele peak identified in hME by the sum of the areas of the mutant and wild type allele peaks. In panel B, the survival of patients in four groups are compared: Low or Int-1 IPSS risk group and *TET2* mutation present (black line), Low or Int-1 IPSS risk group and no *TET2* mutation (green line), Int-2 or High IPSS risk group and *TET2* mutation present (red line), and Int-2 or High IPSS risk group and no *TET2* mutation (blue line). Pairwise comparisons show no significant differences based on *TET2* mutation status in either the lower or higher risk groups.

A.



B.



Supplementary Figure 7: Mutant Allele Frequencies in Individual Samples. Mutant allele frequencies in individual samples with two or more mutated genes validated by quantitative mass spectroscopic genotyping are shown. This includes most co-mutated cases of *TET2*, *RUNX1*, *TP53*, *CBL*, *PTEN*, *NRAS*, *KRAS*, *JAK2*, and *NPM1*. The area of each colored circle indicates the allele frequency of the given mutation. The text box to the left of the circles lists the frequency and nature of each mutation in order of decreasing allele frequency. The FAB class and karyotype for each sample is shown to the right of the circles (NL = normal karyotype, Cpx = complex, +8 = Trisomy 8, Other = other abnormalities not explicitly included in the IPSS). No mutated gene is consistently present with the greatest allele frequency.

